

58-34 SEARCH REQUEST FORM

Attach Paper #11

Requestor's Name: Brownen Loop Serial Number: 09/715,249

Date: _____ Phone: 605-1197 Art Unit: 1636

MAILBOX: 11E12 OFFICE: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

Please search the commercial nucleotide
databases for ~~1-72~~ ²⁰³⁴ of
SEQ ID No. 1.

Please also search for ~~1-72~~ ^{nucleotides}
939-2034
and ~~1-72~~ of SEQ ID No. 1 in the
commercial nucleotide databases.

Thank you.

Edward Hart
Technical Info Specialist
STIC / Biotech
CAL 12C14 Tel: 303-9203

STAFF USE ONLY

Date completed: 1/14/02
Searcher: 1/15/02
Terminal time: _____
Elapsed time: _____
CPU time: _____
Total time: _____
Number of Searches: _____
Number of Databases: _____

Search Site
☒ STIC
☒ CM-1
☐ Pre-S
Type of Search
☒ N.A. Sequence
☐ A.A. Sequence
☐ Structure
☐ Bibliographic

Vendors
☐ IG
☐ STN
☐ Dialog
☐ APS
☐ Geninfo
☐ SDC
☒ DARC/Questel
☐ Other

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: January 14, 2002, 19:04:15 ; Search time 2222.52 seconds
(without adjustments)

15097,845 Million cell updates/sec

Title: US-09-715-249-1_COPY_1_2034
 Perfect score: 2034

Sequence: 1 atgcgacccctccggyacg...acatcgttcggaagcgcacg 2034

| Scoring table: | IDENTITY_NUC | Gapcost 1 0 |
|----------------|--------------|-------------|
| Gapcost 10 0 | | |

Searched: 1472140 seqs, 8248589755 residues

1 number of hits satisfying chosen parameters: 2944280

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

| | | |
|-----|----------------|---|
| 1: | gb_ha: | * |
| 2: | gb_hn: | * |
| 3: | gb_in: | * |
| 4: | gb_cm: | * |
| 5: | gb_ov: | * |
| 6: | gb_pat: | * |
| 7: | gb_ph: | * |
| 8: | gb_pl: | * |
| 9: | gb_pr: | * |
| 10: | gb_ro: | * |
| 11: | gb_sts: | * |
| 12: | gb_sy: | * |
| 13: | gb_un: | * |
| 14: | gb_vl: | * |
| 15: | em_ba: | * |
| 16: | em_fun: | * |
| 17: | em_hum: | * |
| 18: | em_in: | * |
| 19: | em_com: | * |
| 20: | em_or: | * |
| 21: | em_ov: | * |
| 22: | em_par: | * |
| 23: | em_ph: | * |
| 24: | em_pl: | * |
| 25: | em_ro: | * |
| 26: | em_sts: | * |
| 27: | em_sy: | * |
| 28: | em_un: | * |
| 29: | em_vl: | * |
| 30: | em_hqoq_hum: | * |
| 31: | em_hqoq_in: | * |
| 32: | em_hqoq_rod: | * |
| 33: | em_hqoq_hum: | * |
| 34: | em_hqoq_inv: | * |
| 35: | em_hqoq_rod: | * |
| 36: | em_hqoq_other: | * |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|------------|---------------------|
| 1 | 2034 | 100.0 | 3633 | 6 | AX179384 | AX179384 Sequence |
| 2 | 2034 | 100.0 | 5532 | 6 | AR086089 | AR086089 Sequence |
| 3 | 2034 | 100.0 | 5532 | 9 | HMEGFRE | X00588 Human mRNA |
| 4 | 1918 | 94.3 | 2643 | 9 | HUMGFERS | K03193 Human abert |
| 5 | 1876.8 | 92.3 | 2864 | 9 | AF125253 | AF125253 Homo sapl |
| 6 | 1606.4 | 79.0 | 2400 | 9 | HSECF10 | X00663 Human mRNA |
| 7 | 1484.2 | 73.0 | 4188 | 10 | HMNGFR | 7278987 M.musculus |
| 8 | 1484.2 | 73.0 | 5935 | 10 | AE275367 | AE275367 Mus muscu |
| 9 | 1484 | 73.0 | 4194 | 10 | RATNGFR | M37394 Rattus norv |
| 10 | 1481 | 72.8 | 2301 | 10 | HMNGFR | X56698 M.musculus |
| 11 | 1481 | 72.8 | 3704 | 10 | MG03425 | U03425 Mus musculus |
| 12 | 1416.8 | 65.7 | 2618 | 10 | AF124513 | AF124513 Mus muscu |
| 13 | 1207.8 | 59.4 | 1593 | 9 | HSU48722 | U48722 Human epide |
| 14 | 1207.8 | 55.4 | 1868 | 6 | R180039 | 180039 Sequence 1 |
| 15 | 1207.8 | 59.4 | 1868 | 6 | HS095089 | U95089 Human tronc |
| 16 | 1013.2 | 49.8 | 2400 | 5 | CHKEGFR | M20386 Chicken ep |
| 17 | 992 | 48.8 | 1958 | 10 | AF187818 | AF187818 Rattus no |
| 18 | 764.2 | 37.6 | 2243 | 5 | CHKEGF | M7637 Gallus gall |
| 19 | 563.2 | 27.7 | 5766 | 5 | XU053471 | U53471 Xiphophorus |
| 20 | 552.4 | 27.2 | 4138 | 5 | XSMTK | X16891 Xiphophorus |
| 21 | 450.4 | 22.1 | 4060 | 10 | AF041838 | AF041838 Rattus no |
| 22 | 431 | 21.2 | 5484 | 9 | HUMREYKIN | L07868 Homo sapien |
| 23 | 431 | 21.2 | 5501 | 6 | AR040740 | AR040740 Sequence |
| 24 | 431 | 21.2 | 5555 | 6 | AR040741 | AR040741 Sequence |
| 25 | 419 | 20.6 | 3768 | 6 | AX060704 | AX060704 Sequence |
| 26 | 419 | 20.6 | 4530 | 6 | I21124 | I21124 Sequence 9 |
| 27 | 419 | 20.6 | 4530 | 6 | I59745 | I59745 Sequence 9 |
| 28 | 419 | 20.6 | 4530 | 6 | HUMHERA | M11730 Human Lyrus |
| 29 | 419 | 20.6 | 9274 | 6 | AX060703 | AX060703 Sequence |
| 30 | 417.4 | 20.5 | 3768 | 6 | AR034479 | AR034479 Sequence |
| 31 | 417.4 | 20.5 | 4473 | 6 | AR080259 | AR080259 Sequence |
| 32 | 417.4 | 20.5 | 4473 | 6 | HSEB2R | X03363 Human c-erb |
| 33 | 410.6 | 20.2 | 4424 | 10 | RND29339 | U29339 Rattus norv |
| 34 | 409.4 | 20.1 | 4879 | 9 | HUMGFEFBB3 | M29366 Human epide |
| 35 | 409.4 | 20.1 | 4879 | 11 | G28555 | G28555 human SRS S |
| 36 | 409.4 | 20.1 | 4905 | 6 | AR047857 | AR047857 Sequence |
| 37 | 409.4 | 20.1 | 4905 | 6 | I16885 | I16885 Sequence 3 |
| 38 | 409.4 | 20.1 | 4975 | 6 | HUMHERA | M34309 Human epide |
| 39 | 405 | 19.9 | 4062 | 10 | HAMNEU | D16295 Syrian goldf |
| 40 | 402.2 | 19.8 | 3955 | 6 | I21129 | I21129 Sequence 14 |
| 41 | 402.2 | 19.8 | 3955 | 6 | I59750 | I59750 Sequence 14 |
| 42 | 400.6 | 19.7 | 3955 | 10 | RNNEUR | X03362 Rat mRNA fcd |
| 43 | 400.6 | 19.7 | 2385 | 6 | AR082744 | AR082744 Sequence |
| 44 | 400.6 | 19.7 | 2385 | 6 | AR093963 | AR093963 Sequence |
| 45 | 400.6 | 19.7 | 2385 | 6 | AR143949 | AR143949 Sequence |

ORIGIN

| | | | | |
|-----------------------|--------------|--------------|------------|--------------|
| Query Match | 100.0%; | Score 2034; | DB 6; | length 3633; |
| Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| Matches 2034; | Conservative | 0; | Mismatches | 0; |
| | | | Indels | 0; |
| | | | Gaps | 0 |

| | | | |
|----|-----|--|------|
| Qy | 1 | atgcgacccctccgggaagccggggaaagcgtctccctggagcgtgcgtgcgtgcgtctctggccg | 60 |
| Db | 1 | atgcgacccctccgggaagccggggaaagcgtctccctggagcgtgcgtgcgtgcgtctctggccg | 60 |
| Qy | 61 | gcgagtcctgggctctggaggaaaagaattctgcgaagcagcagatlaaagaactcagcag | 120 |
| Db | 61 | gcgagtcctgggctctggaggaaaagaattctgcgaagcagcagatlaaagaactcagcag | 120 |
| Qy | 121 | ctggagcactcttggaaagcactctcttcagagctcccaagagatgtctcaactctggagctg | 180 |
| Db | 121 | ctggagcactcttggaaagcactctcttcagagctcccaagagatgtctcaactctggagctg | 180 |
| Qy | 181 | gtcccttgggaattctggaaaataactactatgctcagaagaaattatgaactcttccttaag | 240 |
| Db | 181 | gtcccttgggaattctggaaaataactactatgctcagaagaaattatgaactcttccttaag | 240 |
| Qy | 241 | accatccagagagctggcctgctcgtctcgtctccatctggccctcaacaagctggagcgaattcc | 300 |
| Db | 241 | accatccagagagctggcctgctcgtctcgtctccatctggccctcaacaagctggagcgaattcc | 300 |
| Qy | 301 | ctggaaaaccctgcagatcaatagaggaataatgtatctagaaaataattctctatgccttagca | 360 |
| Db | 301 | ctggaaaaccctgcagatcaatagaggaataatgtatctagaaaataattctctatgccttagca | 360 |
| Qy | 361 | gtcttactcaactatgatgtcaaatataaaaccggaacttgaagagcttgcctcatgagaattta | 420 |
| Db | 361 | gtcttactcaactatgatgtcaaatataaaaccggaacttgaagagcttgcctcatgagaattta | 420 |
| Qy | 421 | caggaaaaccctgcagagcgctgagctctctagaaaagaaccctgcctctgcaagctggag | 480 |
| Db | 421 | caggaaaaccctgcagagcgctgagctctctagaaaagaaccctgcctctgcaagctggag | 480 |
| Qy | 481 | agcatctcagctggcggagacatagtcagcagctgacctctctcagcaaatatgtcagatgagcttc | 540 |
| Db | 481 | agcatctcagctggcggagacatagtcagcagctgacctctctcagcaaatatgtcagatgagcttc | 540 |
| Qy | 541 | agcattccactggcggagacatagtcagcagctgacctctctcagcaaatatgtcagatgagcttc | 600 |
| Db | 541 | agcattccactggcggagacatagtcagcagctgacctctctcagcaaatatgtcagatgagcttc | 600 |
| Qy | 601 | gctgcagagagaggaactctgcagaagaacttgcaccaaatatcgtgtgccagcagctgctcc | 660 |
| Db | 601 | gctgcagagagaggaactctgcagaagaacttgcaccaaatatcgtgtgccagcagctgctcc | 660 |
| Qy | 661 | ggagcgtctgcgtctggagaaagctcccccagctgacctgctcccaaccagctgtgctcagagctgc | 720 |
| Db | 661 | ggagcgtctgcgtctggagaaagctcccccagctgacctgctcccaaccagctgtgctcagagctgc | 720 |
| Qy | 721 | acagggccccggggagagctgacctgctgctgctcgcgcaaatctcgagacgaagccagctgac | 780 |
| Db | 721 | acagggccccggggagagctgacctgctgctgctcgcgcaaatctcgagacgaagccagctgac | 780 |
| Qy | 781 | aaggaacacctgcgcccccaactcatalgtctctcaaaccccaaccagctgacagatgtgactgac | 840 |
| Db | 781 | aaggaacacctgcgcccccaactcatalgtctctcaaaccccaaccagctgacagatgtgactgac | 840 |
| Qy | 841 | cccgaggagcgaatatcagctcttggctgccaccctgcgtggaagaagctgtcccgtaatatgtg | 900 |
| Db | 841 | cccgaggagcgaatatcagctcttggctgccaccctgcgtggaagaagctgtcccgtaatatgtg | 900 |
| Qy | 901 | gtgaaacaga tcaacagctcgtcgtctccagacgtctggggccggaacagctatgtgaatgtgagga | 960 |
| Db | 901 | gtgaaacaga tcaacagctcgtcgtctccagacgtctggggccggaacagctatgtgaatgtgagga | 960 |
| Qy | 961 | gaagcgctccgcaagctttaaagaagctgcgaagggcctctgcgcaagctgtgttaaaggataa | 1020 |
| Db | 961 | gaagcgctccgcaagctttaaagaagctgcgaagggcctctgcgcaagctgtgttaaaggataa | 1020 |

| | | | | | |
|----------|------|---|---------|--|------|
| Db | 961 | GA | GGGCGTC | CCGAAGTGAACAAAGTGGAAAGGGGCTTCCCGCAAACTGCTAACGGAA | 1020 |
| Qy | 1021 | ggtatctgtgtgaatcttaaaagctcaactctccataaatgtacgaatactaaacaattcaaa | 1080 | | |
| Db | 1021 | GGTATTGTGTGAATTTAAAGACTCACTCTCATAAATGTCACGAATTTAAACACTTCAAA | 1080 | | |
| Qy | 1081 | aactgtcacctccatacaatgtgtgtgtgaactctccataactctgtccgttgagcaetltaaggtgtgaacc | 1140 | | |
| Db | 1081 | AACCTCACCTCCATCAACAGTGGGAGATCTCCATCATCTCTCCGGTGGCAATTTAAGGGGTCACTCC | 1140 | | |
| Qy | 1141 | ttcaacaactactctctctctctgtgatcccaaggaactgtgatalctctgaaaaacgttaagaa | 1200 | | |
| Db | 1141 | TTTCACAACTACTCTCTCTCTGTGATTCACAGCTTGGGCTGAAACAGAGCGGAGCTCATCTCTT | 1200 | | |
| Qy | 1201 | atcacaggggtcttctgtcatltaaggtctgtgcccgtgaataacbaagacaggaacctccatgctctt | 1260 | | |
| Db | 1201 | ATCACAGGGTCTTGTGTGATTCAGAGCTTGGGCTGAAACAGAGCGGAGCTCATCTCTT | 1260 | | |
| Qy | 1261 | ggagaaacctgaagaatacaacggggagaggaaccaagaacaatgtgtcagtcttctctctgtgac | 1320 | | |
| Db | 1261 | GAGAAACCTGAATAATCATACGGGAGGAGCAACAAACATGTGTAAGTTCTCTTTCACATC | 1320 | | |
| Qy | 1321 | gtcagcctgtgaacataaacactcttggatltacgctctccctcaaggaataagtgatgtgaagt | 1380 | | |
| Db | 1321 | GTCACCCCTGGAACATAAACATCTCTTGGGATATTCGCTCCCTCAAGAGATTAAGTATGAGAT | 1380 | | |
| Qy | 1381 | gtgaatacttcaaggaacaacaataatttgtatgtagaataacataaactcgtgaagaataatg | 1440 | | |
| Db | 1381 | GTGATTAATTTCAAGGAACCAAAATTTGTGCTATGCAAAATACATTAACGTGMAAAACATG | 1440 | | |
| Qy | 1441 | tttgggacctccggtcaggaaaacccaataataagcaacagaggtgaanaacagctgtcaag | 1500 | | |
| Db | 1441 | TTTTGGGACCTCGGCTCGAATAACCAAAATTTATAGCAACAGAGGTGAATAACAGCTGCAMG | 1500 | | |
| Qy | 1501 | ggcaacaggccaggtgtcagatgacctgtgtgtctcccccgaaggtctgtctgtggcccggaagcc | 1560 | | |
| Db | 1501 | GGCAACGGCCAGGCTGTGCAATGCCCTTGTGCTCCCGGAGGCTGTGTGGGCCCGGAGGCC | 1560 | | |
| Qy | 1561 | agggaactgtgtctctctgtgcgggaatgtcaagcaggaagcaggaaatgtcgtgtgaacaagtgtcaag | 1620 | | |
| Db | 1561 | AGGGAACCTCGCTCTCTCTTTCGCGGAATGTGACGGCCAGCGAGGAATGTGTAACAAAGTGCAMG | 1620 | | |
| Qy | 1621 | ctctctgtgaggtgtgagccaaggaagttgtgtgtgaatactgtgagtacatgaatgtgacacca | 1680 | | |
| Db | 1621 | CTTCTGTGAGGGGTGAGCCCAAGGAGTTTGTGTGAGAACTCTGATGTGATACAGTGCACCCCA | 1680 | | |
| Qy | 1681 | gaatgtcctgtcctcaaggtccatgtgaacatacactgtcaaggaacggggagccagaacaatgtatc | 1740 | | |
| Db | 1681 | GAGTGTCCCTGTCTCAGGGCATGAACATACCTGTGCACAGGACGGGGACAGCAACATGTATC | 1740 | | |
| Qy | 1741 | cagttgtgcacactcatctgaacggcccccaactgtgttaagaacactgtgcagcaggaagatcag | 1800 | | |
| Db | 1741 | CAGTGTGCCCATCAATTGAACGGGCCCCCACTGCTCAAACTGTGCCGGCAGAGATCATG | 1800 | | |
| Qy | 1801 | ggagaaaaaacaacccccgtctgtgtgaagtacgacagatcgcgcgacatgtgtgccactgtgc | 1860 | | |
| Db | 1801 | GGAGAAAAACAACACCTCGTGTGTGAAGTACGCAAGACCGGGCAATGTGTGCACCTGTGC | 1860 | | |
| Qy | 1861 | catccaactctcacactcctcagatgtcacctgtggccaggtctctgtgaaggtgtgtccaaagaaatg | 1920 | | |
| Db | 1861 | CATCCAAACTCACACTCATCGATGCACACTGGGCCAAGGCTTGTGAAGGCTGTTCACAACGAATGGG | 1920 | | |
| Qy | 1921 | ccctaagaatccggtccatctgcacactgtgaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt | 1980 | | |
| Db | 1921 | CCCTAAGATCCCGTCCATCTGCACACTGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 1980 | | |
| Qy | 1981 | ggccctgtgtgtatcggcctctctcatctgtcgaagagcgcccaacatgtgttcggaagcgtcagc | 2034 | | |
| Db | 1981 | GGCCGTGTGTGTATCGGCTCTCTCATGTGCAGCAAGCGCCCAATGTGTTCGGAACGGCAGC 2034 | | | |
| RESULT 2 | | | | | |
| AR08b089 | | | | | |

[illegible]

```

|||||
Db 547 GTCTATCTACTATGATGCAATTAATAACCGACTGAAGAGAGTCCCATAGATAATTTA 606
Oy 421 cagaaatctctgattggcgccgtgaggttctcagcaaacctctgcccgttcaagctgag 480
Db 607 CAGGAATTCCTGCAATGGCCGGCTGCGTTTCAAGAACAAACCTCCCTGTGTGAACGTGAG 666
Oy 481 agcatcagtcgagggagacatlagtcagcagtgacattctcagcaacatgtcgtgacttc 540
Db 667 AGCATTCAGTGGCGGAGACATAGTCAACATGACTTTCACACCAACATGTCATGTGACTTC 726
Oy 541 cagaaacctctgagcagctcagcaaaagtgtgattccaaagctctccaatgagagctgttcg 600
Db 727 CAGAACCACTGGGCAAGCTCCAAAGTGTGATCCAACTGTGCCCAATGGGAGCTGCTCG 786
Oy 601 ggtcagagagagagagacagtcagcaaaactgacaaatcatctgtgtccagcagtgcttc 660
Db 787 GGTCCAGAGAGAGGAGAACTGCCAGAAATGACCAAAATCATCTGTGCCCACAGCTGCTCC 846
Oy 661 gggcgtctgcccgtgcaagtcgcccaagtgacgtctgcccacaaacagtgctgagcgttc 720
Db 847 GGGGCTGCGGTGCAAGTCCCCAGTGAAGTCTGCGCAAAACCAAGTGTGTGAGGCTGC 906
Oy 721 acaagccccgggagagagcagctgtctgtctgcgcaaatctccagagagagagcagctgc 780
Db 907 ACAGGCCCCCGGAGAGGAGAGTCTGTGTCTCCGAAATTCGAGACGAGACCAAGTGC 966
Oy 781 aagcaacacgtgccccacacacacacacacacacacacacacacacacacacacacacac 840
Db 967 AAGGACACTGCCCCACATCATGTCTACAAACCCACACACATGACATGATGATGATGATG 1026
Oy 841 cccgagagagagagagagagagagagagagagagagagagagagagagagagagagagag 900
Db 1027 CCGGAGGCAAAATACACTTTGGTGTGACCTGTGCAAGAGAGTGTGCTCCGTAATTTATG 1086
Oy 901 gtgacagacacagcagcgtctgtctgtcagcagcgtctgtgagagagagagagagagag 960
Db 1087 GTGACAGATCACGGCTGTGTCTGCTCCAGGCTGTGGGGCCGACAGGTATGAGATGAGAG 1146
Oy 961 gacgagcgtccgcaagtgtaagagtgagagagagagagagagagagagagagagagag 1020
Db 1147 GACGGCGCTCGCAAGTGTAAAGAGTGCAGAAAGGCTGTGCGGAAAGTGTGTAAGAGGATA 1206
Oy 1021 ggtatgtgtgtaatttaagacacacacacacacacacacacacacacacacacacacac 1080
Db 1207 GGTATGTGTGTAATTAAAGACTCACTCTCCATTAATGCTAGCAATTAATTAACACTTCAA 1266
Oy 1081 aactgacacctcaatcagtgagacacacacacacacacacacacacacacacacacacac 1140
Db 1267 AACTGACACTCTCATCAGTGGGAGTCTCCACATCTCTCGGTGSCATTTAGGGGTGACTCC 1326
Oy 1141 ttcacacatctctctctctctctctctctctctctctctctctctctctctctctctct 1200
Db 1327 TTTCACACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1366
Oy 1201 atcaacaggtttctctctctctctctctctctctctctctctctctctctctctctct 1260
Db 1387 ATCAACAGGTTTGTGCTGATTCAGGCTTGGCTGAAACAGAGACGACCTCCATCTCTTT 1446
Oy 1261 gggaaacctagaaatcatagcggagagagagagagagagagagagagagagagagagag 1320
Db 1447 GAGAACTAGAAATCATACGGGAGAGACCAACACATGTCAGTCTTTCTTTCAGATG 1506
Oy 1321 gtcagcctgaaataacatctctgtggaattagcgtccctcaagagagagagagagagat 1380
Db 1507 GTACACCTGAAACATTAACCTCTTGGGATTCCTGCTCCCTCAAGAGATTAAGTATGAGAT 1566
Oy 1381 gtgataattcaggaacaaatctgtctatgcaatacaataaactgtaaaactg 1440
Db 1567 GTGATTAATTTCAGAAACAAATTTTGTGCTATGCAAAATTAACATGGAATAAATCTG 1626
Oy 1441 ttgtggacctcgtctcagaaacaaatattatagcaacagaggtgaaacagctgcag 1500
|||||

```

```

Db 1627 TTGGGACCTCCGCTGCAAGAAACAAATTTATAGCAACAGAGTGAATAACACTGCAAG 1686
Oy 1501 gccaagagcaggtctgcaatgctgtgtctcctccagagagctgtgagccgagagcc 1560
Db 1687 GCAACAGCCAGGCTCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1746
Oy 1561 aggaactgctctctctctctctctctctctctctctctctctctctctctctctctct 1620
Db 1747 AGGACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1806
Oy 1621 ctctgagaggttagcagcaagagagagagagagagagagagagagagagagagagagag 1680
Db 1807 CTCTGAGAGGCTGAGGCAAGGAGGAGTTTGAGAGAACTGTGATCATACAGTCCACCA 1866
Oy 1681 gagtgcctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1740
Db 1867 GAGTGCCTGCTGAGGCAATGAGCAATGAGCAATGAGCAATGAGCAATGAGCAATGAGCAAT 1926
Oy 1741 cagtggtcccaacacacacacacacacacacacacacacacacacacacacacacacac 1800
Db 1927 CAGTGTCCCACTACATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1986
Oy 1801 ggaagaaacacacacacacacacacacacacacacacacacacacacacacacacacac 1860
Db 1967 GGAAGAAACACACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2046
Oy 1861 catcaaacacacacacacacacacacacacacacacacacacacacacacacacacacac 1920
Db 2047 CATCAAACTGCAACCTGAGATGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2106
Oy 1921 cctaagatccgcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1980
Db 2107 CCTAAGATCCGCTGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2166
Oy 1961 ggccttgagagagagagagagagagagagagagagagagagagagagagagagagag 2034
Db 2167 GCCCTGGAGATGGGCTCTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220

```

RESULT 4
HUMEGFRS
LOCUS
DEFINITION Human aberrant (short) epidermal growth factor receptor mRNA, complete cds.
ACCESSION K03193.1 GI:181984
VERSION K03193.1
KEYWORDS epidermal growth factor; epidermal growth factor receptor; erbB oncogene.
SOURCE Human epidermoid carcinoma cell line A431, cDNA to mRNA, clone pE15.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2643)
AUTHORS Merlino, G.T., Ishii, S., Whang-Peng, J., Knutsen, T., Xu, Y.-H., Clark, A.J., Stratton, R.D., Wilson, R.K., Ma, D.-P.P., Roe, B.A., Hunts, J.H., Shimizu, N., and Pastan, I.
TITLE Structure and localization of genes encoding aberrant and normal epidermal growth factor receptor RNAs from A431 human carcinoma cells.
JOURNAL Mol. Cell. Biol. 5 (7), 1722-1734 (1985)
MELINE 85267689
COMMENT A draft entry and printed copy of this sequence were kindly provided by G. Merlino (07-OCT-1985).
FEATURES
source
 1..2643
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="7p13-p12"
 <1..>2643
 /gene="EGFR"
 /note="EGFR(S) mRNA: G00-120-610"
 1..2643
 gene

probably also shared

[illegible]

| DB | 1986 | CA01GTCACACTACATATGAGS6GCCGCCACATGCGCAAGACACTGCGCGGACAGAGTCAATG | 2045 |
|----------------------------|---|---|----------------------|
| Qy | 1801 | ggagaaaacaaacaccctggtcttggaaagtcagcagacgcggccatgtctgtccaccctgtc | 1860 |
| Db | 2046 | GGGAAAAACACACACCCCTGTCTGTGGAAAGTACGACAGCCGCGCCCATGTGTGCCACTCTGTGC | 2105 |
| Qy | 1861 | catccaactgacactacacg | 1880 |
| Db | 2106 | CATCCAAACTGCACCTCAACG | 2125 |
| RESULT | 6 | | |
| LOCUS | HSEG01 | 2400 bp | mRNA |
| DEFINITION | Human mRNA fragment for epidermal growth factor (EGF) receptor. | | |
| ACCESSION | X00663 | | |
| VERSION | X00663.1 | GI:311112 | |
| KEYWORDS | cellular oncogene; epidermal growth factor receptor; erbB cellular oncogene; oncogene. | | |
| SOURCE | human. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | |
| AUTHORS | 1 (bases 1 to 2393) | | |
| TITLE | Xu,Y.H., Ishii,S., Clark,A.J., Sullivan,M., Wilson,R.K., Ma,D.P., Roe,B.A., Mellino,G.T. and Pastan,I. | | |
| JOURNAL | Human epidermal growth factor receptor cDNA is homologous to a variety of RNAs overproduced in A431 carcinoma cells | | |
| MEDLINE | Nature 309 (1991), 806-810 (1984) | | |
| REFERENCE | 2 (bases 1 to 2400) | | |
| AUTHORS | Merlino,G. | | |
| TITLE | Direct Submision | | |
| COMMENT | Submitted (28-MAY-1985) to the EMBL/GenBank/DBJ databases | | |
| FEATURES | Data kindly reviewed (28-05-1985) by Merlino G. | | |
| Source | Location/Qualifiers | | |
| | 1..2400 | | |
| | !..organism="Homo sapiens" | | |
| | Eukaryota; taxon:9606 | | |
| | <1..->2400 | | |
| | /note="EGF (1 is 2nd base in codon)" | | |
| | /codon_start=1 | | |
| | /protein_id="CAA25282.1" | | |
| | /db_xref="GI:4378981" | | |
| | /db_xref="SWISS-PROT:P00533" | | |
| | /translation="ILHGAVRFSNPNALCNVESIQMDRIYSDPFLSNMSMPDNHNLGSGCQDSCPNSSGACGEGNOKLTKITLQAGGSGRGRGSPDCHCNCAAGCGPRPESQICAKRREARCTKCTGCPMLYLNPTTMYMDNNGYSPGATCYKCPRIYVYVYTHDSCVRACGADSTYMERDGYRKKCGDEPKRCYVNGIGERDSSLINATIRIKERKNCSTISDILLPLVAFRSDSTPHPLADPOEIDLITKYKELTFELLIQMPEPNRTDILHAFMEHLIGRTKQHGQESLAVVSLITSILGRSLIKELISQDVI1SGKNALCYANNTINWKRLPFSIQOKTKR1ISNKGESCAKGGVQALGSLSPGCGPRPACVSRNWSGREGECVGRKLLLEGEPREFVENSBCIQCEPCILQAMNLTCTGGPNQCTQCAHYIDGPHCVCTCPACGMEENNTLWVKYADAGVCHLCHPNTCYGCGPGLGEGCPNGKPIESTATGCMKATGALLLVVALGIGLFWRBRHIVRKRTLRLIDRELELVPPLPSEAPNOLAURLIKLETERRKIKVLGSGAGTGVYGLMIPGEVKVLPVALIKELREATSPKANKETLDEAYVWASVDNHCVRGLGCTISTYQVLTQIMLMPGCLIDVREHNDISQYLLNNCYQIAGCMYNIIEPRRLYHBDIAARVNLVYKTPQVYKLTDPGLAKLIGAEKREYIABSGYPIYNNMALLESILRITTHOSDWSYSGVYVMEITPFSKRPDIQIPASEISITLKEGERLPQPICTID" | | |
| BASE COUNT | 615 a | 655 c | 659 g 471 t |
| ORIGIN | | | |
| Query Match | 79.0%: | Score 1606.4: | DB 9; Length 2400; |
| Best local Similarity | 99.9%: | Pred. No. 0; | |
| Matches 1607; Conservative | 0; | Mismatches | 1; Indels 0; Gaps 0; |
| Qy | 427 | atccgcgatgagcgcgttgcgttcagcaacaacccctgcctctgtcaacgtygagagcacc | 486 |
| Db | 1 | ATCCTTCATGCGCCGCGGTCGCTTCAACAAACCCCTCCCTGCGAACGTGGAAGACATC | 60 |
| Qy | 487 | caatggcggagacatagtcagcagtagtcttcctgcagcaaatctgatagtgacatctccagac | 546 |

| | | | |
|----|------|---|------|
| Dy | 61 | CATGGCGGAGATATGTACACTGACTTTCACAGCAAACTGTTCATGTGACTTTCAGAAC | 120 |
| Oy | 547 | caacttggagcagcttcccaaaagtgtgtatcaagctgttccaaatggagctgtcctgggtgtca | 606 |
| Dy | 121 | CACCTGGGAGCTGTCCAAAATGTGTATTCACAGCTGTCCCAATGAGAGCTGGGGTCCA | 180 |
| Oy | 607 | ggagagagaaacttgcagaaacttgaccaaaatcaatcttgcagcagatgtctccgggagc | 666 |
| Dy | 181 | GGAGAGGAGAACTGCCAGAAACAGACAAAATCATCTGTGCCAGAGAGCTCCGGGGC | 240 |
| Oy | 667 | tgcgtgtgcaagtcoccccaatgtactgtctccaaaccaagatgtctcagctgtccagagc | 726 |
| Dy | 241 | TGCCGTGGCAATGCCCACTGACTGTCTGCCAACCAGTGTGCTGACGTGCACAGGC | 300 |
| Oy | 727 | ccccggagagcagactgtcctgtgtctgcgcgaaatttcgaaagaaagccacgtgtgaagac | 786 |
| Dy | 301 | CCCCGGAGAGCACTGCTGCTGTCTCCGCAAAATTCCAGACCAAGCCACGTGCAGAGAC | 360 |
| Oy | 787 | acttgcoccccaactcaatgtctctcaaaccccaaccacgtatccaagatgtgtgaaccgag | 846 |
| Dy | 361 | ACCTGTCCCCCACTCATGTACTCTGTACAAACCCCAACAGTACCAAGATGTGAAGTGAACCCGAG | 420 |
| Oy | 847 | ggcaaatagaagtttgtgtgcacactgtcgttgaagaagtgtcccgtaactatgttgtgtga | 906 |
| Dy | 421 | GGCAAAATACAGTTTGGTGTGCCACTCGCGAAGAAAGTGTCCCGTAATTATGTGGTGTACA | 480 |
| Oy | 907 | gaccaaggtctgtgcgtccagaccccttgaggcgcaagctatagaattgaagaaagaagc | 966 |
| Dy | 481 | GATCACGGCTGTGTGTCCAGCCTGTGGGGCGACACTGTAGATGGAGGAGAACGGC | 540 |
| Oy | 967 | gtccgaagaattgaagaagtgtcgaaggccttgcgcgaaagtgtgtgaagaaatgaatt | 1026 |
| Dy | 541 | GTCCGCAAGTGTAAAGACTGCAGAGGGCTTGTCCCAAGGTGTAAACGAATTAAGTATT | 600 |
| Oy | 1027 | ggtgtgaattaaagactcaactctccataaattgtctacgaatatlaacaactcaaaaatgac | 1086 |
| Dy | 601 | GGGAAATTTAAAGACTCACTCTCCATAAATGTCTAGAAATATTAAACCTTCAAAAATGTGC | 660 |
| Oy | 1087 | acctcatcagatgtgcgatctccacaactctgtccggtgtgcatlttaggtgtgaccccttcaa | 1146 |
| Dy | 661 | ACCTCATCTAGTGGGCAATCTCCACATCTCCGGGGGATTATGAGGGGTACACTCTTACAA | 720 |
| Oy | 1147 | caatacctccctctctgtgataccaagagaactgtatcttgcgaaacgttgaagaaatactca | 1206 |
| Dy | 721 | CATACCTCCCTCCGTGATTCACAGGAACGTGATTTCTGAAAACCGTAAAGGAATGTACA | 780 |
| Oy | 1207 | ggattttgtgatttcagagcttggccttgaaaacagaaagagccctcaatgcttgaagac | 1266 |
| Dy | 781 | GGGTTTTGTGTATTCACAGGCTTGGCTGTGAAAACAGAGACCTCCATGCTTTGAGAAC | 840 |
| Oy | 1267 | ctbagaatcatacggcgagagaccaagaacaatgtgtcaagtttctctcttgcagtgctcagc | 1326 |
| Dy | 841 | CTTGAATATATACGGGGAGAGACCAAGCAACATGTGTCAAGTTTCTCTTGGAGTGTCAAGC | 900 |
| Oy | 1327 | ctbagaatacaactccttggagtaatgaagctccctcaagagagataagtgatggagatgtgata | 1386 |
| Dy | 901 | CTTGAATCATTAACCTTGGGATTTACCTCTCCCTCAAGAGATTAATGATGTGAATGTGATA | 960 |
| Oy | 1387 | attccagaaacaaaatttgtgtctatgcgaatatacaataaaacttgcgaaaaactgttttgg | 1446 |
| Dy | 961 | ATTTTCAGCAAAATTTTGTCTGTATTCGAATATCATATAAATCTGAAAAACGTGTTGGG | 1020 |
| Oy | 1447 | acctccgtgcagaaacaaataattataagcaacagatgtgaaaacagctgtcagaagccaca | 1506 |
| Dy | 1021 | ACCTCGGTGCATAAACCAAAATTTATTAACATCAAGAGGTGAAAAAGCTGTGCAGAGCCACA | 1080 |
| Oy | 1507 | ggccagatctgtcattgtccttgtgtctcccccgaggctgtcttgaggcccggaagcccaaggac | 1566 |
| Dy | 1081 | GGCCAGAGTGTGCATGGCTTGTGTCTCTCCCAAGAGGCTCTGTGGGCTCCGGAGCCAGAGAC | 1140 |
| Oy | 1567 | tgtgtctcttgcgcgaatgtcaacgcgaagcagaagatatgtgtgcgaagtgtcagactgtt | 1626 |

| DB | 1141 | TGATCTCTGCGGAGATGTCAGCCGAGGAGGAGGAAATGCGTGAACAAATGCACACTTTCG | 1200 |
|------------|---|---|-------------|
| QY | 1627 | gagggatgagcgaaggaaattgttgaaactctgagtgtataacgtgccacccaagatg | 1686 |
| Db | 1201 | GAGGGATGAGCGAAGGAGATGTTGTGAAACTCTGAATGTGATACATGTGCACCAAGATGC | 1260 |
| QY | 1667 | ctgccttaagcaatgaacatacactgtcacagaagcaggagacacaaatgataccagt | 1746 |
| Db | 1261 | CTGCGCTCAGCCCATGAAATCACTCACTGCACAGGACGGGACACGACAACTGATTCACAGTGT | 1320 |
| QY | 1747 | gcccaactaactgaagcggcccaactctgcttaaaagacctgcggcagagatcatgggagaa | 1806 |
| Db | 1321 | GCCCACTACTGATGAGCGGCCCCCACTGCGTCAAAAGCACTGCCCGGAGGATCATGAGAGAA | 1380 |
| QY | 1807 | aacaacacccctggtcttgaaatgaacagagcggcgacatgtgtgcacactgtgcata | 1866 |
| Db | 1381 | AACAACACCCCTGTGCTGGAATGACGACAGCCCGCACATGTGTGCACACTGTGTGCATCCA | 1440 |
| QY | 1867 | aactgcacctaaggaatgaactcgtggcccaagtccttgaaagctgtccaaagatgggactaa | 1926 |
| Db | 1441 | AACGCACTCAGGATGACGATCGGCGACAGTCTTGAAAGCTGTGCACAAATGAGGCTAAG | 1500 |
| QY | 1927 | atcccgatcaatgcacactgggaatggtgaggggccctccctctgtcgtgtgtgtgtgcctt | 1986 |
| Db | 1501 | ATCCCGATCAATGCACACTGGGATGATGATGGGGCCCTCTTCTCTGCTGTGTGTGCGCTTG | 1560 |
| QY | 1987 | ggagtcggccctctcatcgcgaagcgcacacatgcttcggaagcgacg | 2034 |
| Db | 1561 | GGAATCGGCTCTTCTCATGCGAAGGCGCCACATCGTTGCGGAACGCGACG | 1608 |
| RESULT | 7 | | |
| LOCUS | MMBGR | 4188 bp | mRNA |
| DEFINITION | M.musculus (BALB/c) Epidermal Growth Factor Receptor mRNA. | | 17-FEB-1997 |
| ACCESSION | X78987 | | |
| VERSION | X78987.1 | GI:4888830 | |
| KEYWORDS | epidermal growth factor receptor. | | |
| SOURCE | house mouse. | | |
| ORGANISM | Mus musculus | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus | | |
| AUTHORS | 1 (bases 1 to 4188) | | |
| TITLE | Hibbs,M.L. | | |
| JOURNAL | Direct Submission | | |
| REFERENCE | Submitted (05-APR-1994) M.L. Hibbs, Ludwig Institute for Cancer, Research Melbourne Tumour, Biology Branch P O Royal, Melbourne | | |
| AUTHORS | 2 (bases 1 to 4188) | | |
| TITLE | Avivi,A. Skorecki,K., Yavor,A. and Givol,D. | | |
| JOURNAL | Promoter region of the murine fibroblast growth factor receptor 2 | | |
| REFERENCE | (hek/Kcfr) gene | | |
| AUTHORS | oncogene 7 (10), 1957-1962 (1992) | | |
| TITLE | 3 (bases 1 to 4188) | | |
| JOURNAL | Parta,B.C., Das,S.K., Andrews,G.K. and Dey,S.K. | | |
| REFERENCE | Expression of the epidermal growth factor receptor gene is | | |
| AUTHORS | regulated in mouse blastocysts during delayed implantation | | |
| TITLE | Proc. Natl. Acad. Sci. U.S.A. 90 (1), 55-59 (1993) | | |
| JOURNAL | 93126380 | | |
| FEATURES | Location/Qualifiers | | |
| SOURCE | 1..4188 | | |
| | /organism="Mus musculus" | | |
| | /strain="BALB/c" | | |
| | /db_xref="taxon:10090" | | |
| | /chromosome="11" | | |
| | /tissue_type="liver" | | |
| | 60..3692 | | |
| | /codon_start=1 | | |
| | /product="Epidermal growth factor receptor" | | |
| | /protein_id="CAA55587.1" | | |
| | /db_xref="GI:488831" | | |
| | /db_xref="MCD:MG1:95294" | | |
| CDS | | | |

[illegible]

| | | | |
|----|------|--|------|
| Dp | 764 | CAGAGCATCCAGACGATGGTCCCAATGTATTCACAGCTGTCCCAATGGAACCTGTGG | 823 |
| Oy | 601 | ggtgtcggagagagagaaacttcgcaaaaactgacaaaatcattctgtcccgacgtgtcc | 660 |
| Dp | 824 | GCAGCAGCAGACAGACACTGCGCAAAATTGACCAAAATCATCTGTGCCACGAAATTTTC | 883 |
| Oy | 661 | gggtcgtcgtgtgcaagttcccccgtgacgtgtgcacaaacagtggtgtcgaagctgtgc | 720 |
| Dp | 884 | CATCGCTGTGTCGGAGAGGTCCCTCCACGACTGCTCTCCCAACCAATGTGTGTGGGGGT | 943 |
| Oy | 721 | acatggccccggagagagagactgcctgtgactgtgcgcaaatcttcgagacgagagccagtgc | 780 |
| Dp | 944 | ACAGGGCCCCCGAGAGAGTGAATGCTGTGTGCTCTCCAAAATTCCAAAGTAGAGCCACATGC | 1003 |
| Oy | 781 | aaagaaactgtccccccactctgactctacaaaccccaacagatcagaatgtgagtgtgaac | 840 |
| Dp | 1004 | AAAGACACTGTCCGACACACTGATGTGTACAAACCCACACACTGTACAGATGTGTCAAC | 1063 |
| Oy | 841 | cccgaaaggcaaatlaaagtttggtagacactgtcggaagaagtgtcccgtaattatgtgtg | 900 |
| Dp | 1064 | CGTAAAGGGAATTAAGATTTGGTGCGCCACCTGTGTGAAGAAGTCCCCCGAAACTACGG | 1123 |
| Oy | 901 | gtgacagatcaccgctcgtgtgtccgacgtgtggtggccgagactatlgagatgtgagaa | 960 |
| Dp | 1124 | GTACAGCATCAAGGCTCATGTGTCTCGAGGCTGTGGGCTTGACTCTACGAAAGTGAAAG | 1183 |
| Oy | 961 | gacgggtctccagatgtgaagaagtcggaagaagctctgtccgaagtgatgtacacgtgata | 1020 |
| Dp | 1184 | GATGGCATCCCGCAATGTAATAAATGTATGGGCCCTGTCCCAAAAGTTGTATGGCATTA | 1243 |
| Oy | 1021 | ggtatctgtgtgaatttaaaagactcactccataaalgtaagaatataaactctcaaa | 1080 |
| Dp | 1244 | GGCATTTGGTAATTTAAAGACACTCTCCATAAATGTGTACAAACATCAACACTTCAAA | 1303 |
| Oy | 1081 | aaatgaactctcaataagatgtgagatctccacatctctgcgcgtgtgagatctlaaggtgtacc | 1140 |
| Dp | 1304 | TACTGCATGCGCATAGCGGGAGACCTTCAACATCTGTCCAGTGGCTTTAAGGGGATTC | 1363 |
| Oy | 1141 | ttcacacataactctcctcctctgtgatlccacaggaactgtgattctcgtaaacccgtlaaagaa | 1200 |
| Dp | 1364 | TTTACCGCGCACTCTCTCTTATGACCCACGAGAACTTAAGAAATTTCTAATAACCTTAAGGA | 1423 |
| Oy | 1201 | atcacaggttcttgtgatactgagctgtgcctgtgaaaaacagacgagctccatgtgcctt | 1260 |
| Dp | 1424 | ATTAAACGCTTTTGTCTGTATTCAGGCTTGAGGCTGTATTACTGACTGTACCTCCATGCTTTC | 1483 |
| Oy | 1261 | gaabaacctgaataataatcgcgcgcgcgacgagacgaacaaatgtgtccagttctctctgtcgtc | 1320 |
| Dp | 1484 | GAGAACTCTGAATATATATCTGTCGACAAACAAAGAACTGTGTGCTTTCTTGTGGGGTTC | 1543 |
| Oy | 1321 | gtcagaccctgaataataactcctgtggatctlaagctccctcagaagagataaagtatgtgagat | 1380 |
| Dp | 1544 | GTTCGGCTGTACATCATCACTACTGTGGGCTGTCTTCCCTCAAGGAGATCACTATGGGAT | 1603 |
| Oy | 1381 | gtgataatattcaggaataaataattgtgtctatgtcaatataataataactgtgaataaactgt | 1440 |
| Dp | 1604 | GTATATATTTCTGGAAACCCGAATTTGTGTCTGCAAAACAAATAAATGTGGAAAAACTC | 1663 |
| Oy | 1441 | cttgtgacctcgtgtcgtgaataaaccaaatataagcaacagaggtgtgaataacagctgtgcaag | 1500 |
| Dp | 1664 | TTTGGGACACCCCAATCAAGAAAACAAATAATCAATGAACAACAGCTGTGAAGAACTGCAAG | 1723 |
| Oy | 1501 | ggcaacagacaggtgtctgtccatgtcctgtgtctcccccggggctctctgtggtcccgagacc | 1560 |
| Dp | 1724 | GGCTGTAAACACAGTCTGCATATCTTTATGCTCTCTCGAAGGCTCTGTGGGCTCTGAGCC | 1783 |
| Oy | 1561 | agagaaactgtctctctcctcgaaatgttcagccgaagacagagaaatgcgtgtgacaagtgtgag | 1620 |
| Dp | 1784 | AGGGAAGTGTGTCTCTCTCCGCAAAATGTGTACACAAGCAGGAGAGTCTGTGAGAAATGTCAAC | 1843 |
| Oy | 1621 | ctctctgtgaggtgtgagccaaaggaggtctgtgtggtgaactctgtgattcatacagtgaccaca | 1680 |
| Dp | 1844 | ATCTCTGTAGGGGAACCAAGGAGTTTGTGTGAATAATTTGTGATCTATTCACAGTGTCCATCA | 1903 |

[illegible]

Dh 1120 GATGATCCGCAAGTGTAAAAATGATGAGCCCTGTCGCAAACTTTGTAATGCCATA 1179
Oy 1021 gataatgtaatttaagaactcactccatcaatagtacgaatatlaaaccttcaaa 1080
Dh 1180 GGCAATTGTGTAATTAAACACACTCTCCATTAATGTCATCAAAACCACTTCAAA 1239
Oy 1081 aactgacactcactcagtgagatcacaatcctgcgcgttgacattagagggtgacctc 1140
Dh 1240 TACGACCTCCGCAAGCGGGGACCTTCACATCCCTGCGAGGCGCTTAAAGGGGATCT 1299
Oy 1141 ttacacatactcctccttgatgacacagaagatgatatctgaaacccgttaagaa 1200
Dh 1300 TTCACGGCAGCTCTCTCTAGACCCAGAGAACTAGAAATTTAAAAACCGTAAAGAA 1359
Oy 1201 atccagaagtgttgatgatatcaggttgagctgaaacagaagagccctcctcctt 1260
Dh 1360 ATTAACAGGCTTTTGGCTGATGACGCTTGCGCTGATACCTGACCTGACCTGACCTTTC 1419
Oy 1261 gagaaactagaatacaacgagcagcaagcaacacatgagtttctcctgacgtc 1320
Dh 1420 GAGAACCTAGAAATTAATACGTGGGCAACAAAGCAATGCTCAGTTTCTTGGCGCTC 1479
Oy 1321 gtacgactgacataacactccttgatgagatgagctcctcagaagagatgaatgagat 1380
Dh 1480 GTTGGCTGAAACATCACTACTGAGGCTGCTTCCCTCAAGAGATCAGTGAATGAGGAT 1539
Oy 1381 gtgataattcaggaacacaaatctgtgctatgacaaatacaataaactggaanaactg 1440
Dh 1540 GTGATCTATTCTGTAACCAAGCAATTTGCTACGCAAAACAACTGCAAAACCTC 1599
Oy 1441 ttgagaccccgagcaggaacacaaatataaagacagagtgaaacagctgcaag 1500
Dh 1600 TTCGGACACCCATCATGAGAAACCAATCATGACATCAACAGCTGAGAAAGACTGCAAG 1559
Oy 1501 gcaacagcagagctgctgacgtcgtgctcctcccgagagctgctgagggcccgagacc 1560
Dh 1660 GCGGTGAAACAGCTGTGCAATCTTTATGCTCTGGAAGGCTCTGAGGCTGAGCCCT 1719
Oy 1561 aggaagcagctcctcctgagcagatgacagcagagcaaggaatcgtgagcaagtgcag 1620
Dh 1720 AGGAGCTGTGCTGCTGCGCAGATGTGACGACGAGCGAGGTGCGTAGAATGGAAC 1779
Oy 1621 ctctgagaggtgagcagaagagaggttctgagagactctgagatgatacagtgccacca 1680
Dh 1780 ATCTTGAAGGGGAGACCAAGAGATTGTGAAAATTTCTGAATGATCCAGTGCATCCA 1839
Oy 1681 gactgctgctcctcagcagcatalcaatcaccgtgacagagcagggagccagacatgatac 1740
Dh 1840 GAATGTCTGCCGACGAGGCAATGACATCCTGTACAGCGAGGCGCCAGACACTGCAATC 1899
Oy 1741 cagtgctcctcactacatgacagcagccctcctcgtcaagagcctcccgagagagatg 1800
Dh 1900 CAGGTGCTCCCACTACATGATGATGCGCCACACTGTGTCAAGACCTGCGAGCGCATAG 1959
Oy 1801 ggaagaaacacacactgctgagagatgacagcagcagccagccatgtgtcaccctgtgc 1860
Dh 1960 GGAGAGAAACAACTCTGTGCTGGAAGTATGACAGATGCAATTAATGCTGCACCTATGTC 2019
Oy 1861 catcaaacactgacactgagatgcatcagatgagcagagctcttgaagcgtgtgca-----acg 1914
Dh 2020 CAGCCCAACTGATACCTGTGATGTGCTGCTGAGGCGCCAGGTCTTCAAGGATGTGAAGTGTGCGCA 2079
Oy 1915 aatggagctcaagatcccgctcactgacacagatgagatgagtgagggccctcctctgtgctg 1974
Dh 2080 TCTGGGCAAGATACATATTTGCTGCTGAGGATGTGAGGCTGCTCTCTATAGTGTG 2139
Oy 1975 gctgagcctcagggatcagcctcctcactgacagagcagccacatcgttctggaagcgac 2033
Dh 2140 GTGTGTGCGCTTGTGATGCTGCTATTCATGACGAAAGCGTCAATGTTGGAAGCGTAC 2198

RESULT 11
MM003425

LOCUS MM003425 3704 bp mRNA ROD 06-MAY-1994
DEFINITION Mus musculus B6/C3 EGF receptor mRNA, complete cds.
ACCESSION U03425
VERSION U03425.1 GI:458123
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 3704)
AUTHORS Luettke, N.C., Phillips, H.K., Qiu, T.H., Copeland, N.G., Barr, H.S., Jenkins, N.A. and Lee, D.C.
TITLE The mouse waved-2 phenotype results from a point mutation in the EGF receptor tyrosine kinase
JOURNAL Genes Dev. 8, 399-413 (1994)
MEDLINE 94170986
REFERENCE 2 (bases 1 to 3704)
AUTHORS Lee, D.C.
TITLE Direct Submission
JOURNAL Submitted (10-NOV-1993) David C. Lee, Department of Microbiology & Immunology, University of North Carolina at Chapel Hill, UNC Lineberger Comprehensive Cancer Center, Campus Box #7295, Chapel Hill, NC 27599-7295, USA
FEATURES
source location/Qualifiers
1..3704
/organism="Mus musculus"
/strain="B6/C3"
/db_xref="taxon:10090"
/sex="male/female combined"
/tissue_type="liver"
/dev_stage="adult"
51..3683
/codon_start=1
/product="EGF receptor"
/protein_id="AA017899.1"
/db_xref="GI:458124"
/translation="MRPSGTVARFTLLVLLTALCAAGALEEKVKCCGTSNRLDGLFP
EDHFLSLQRMVNCNEVAVLNLITVYORNDLSPKLTQDEAGVYALANLVEPILPE
NLIQIRGALVNTYVALLAILSNVTGNTGRLSPKLTQDEAGVYALANLVEPILPE
TIOMNDIYQNVPMNSMDSQSHSSPCPKDPCSPGSCGSGEENCKLKKITCAAO
CHRPGRSPSDCCHNOCAAGCTSPRESQDINCOKQODATCKDPCMLYNTPTQ
MVNPEGYSFSGCAKCKPRNVYVDHSGCAAGCPDYEEVEEDGIRKCKKCGPCR
KYCNQIGIGERKDLISNATNIRKFKCTAISDHLILVAKCQSFRTPELREL
ELTKVKEITGTFLLIOAMPNDMLAFLELFTIRGTRQKQSPFLAVGNIISGL
RELKLSGSDVITISGNRLICANTIMKKLFGPQNPCKTQIMNRAEKCKRVNHCNP
LCSBSCGPEPDRVSCQNSGREGVEKNTLIEGPPEPENSQCIOCPEDLPDA
MNITGCPDNCIOCAHTIDGPRVCTCPAGIMGRNLTLYKTDANNVCHLHANC
TYGACGPELQCEVWPSGPRTPSLATGIVGGLFIVVALGIGLFRARRHYVRBTJR
RLQORELVEPLTPSGEAPNOAHRLILKEFEKRIYVLSGAGFVYGLMIPGEVY
KIPVAIKELREBATSPPKANKETLDEAYVMAVDNPHVCRLLGICLSTVOLITOLMPG
CLIDVVRHKDNISQVLLNMCVQIAGKNYLEDRLVDRDLAARNVLYKTPQHKIT
DFGLAKLGAEBEKYHAGGKVPTEKMALESILHRTYTHOSDVMAGYVVELMTFGS
KPYDQIPASDITSLIEKGERIPDPICTIDVYATVIMKCMIDADSRPRLIEFSK
MARDDPRLVIOGDERHIFSPDTSNRYRLMDEEMEDVYVADAEVLIPOGFNPS
TSRTLLSLSTNSNTSVACINRNGSRKDEATLQRSSPNAVGNPRLYNTAOPIC
VEEYVNSVPRFAGSVONRYVHNPDLHPRAGDLTYDPHNAVGNPRLYNTAOPIC
LSGSGNPSALMIOKSHOMSLDNPDODFPEKTPKPNGLFPGTAEENETLRVAPPS
SEPTGA"

BASE COUNT 1019 a 961 c 923 g 801 t
ORIGIN

Query Match 72.8%; Score 1481; DB 10; Length 3704;
Best Local Similarity 83.3%; Pred No. 0;
Matches 1698; Conservative 0; Mismatches 335; Indels 6; Gaps 1;

Oy 1 atgcaacctccgagcagcgagggcagcgctcctgagcgtgctgctgctgctgccc 60
Dh 51 ATGCGACCTCTCAGGACCCCGAGAACACACACTGCTGCTGCTGACCGGCTCTCGCC 110
Oy 61 gcaagctcagcgtcaggaagaaagatgctcgaagcagcagtaacacagccacagcg 120
Dh 111 GCAGGTGGGCGTGTGGAGGAGAAAGAGTGTGCGAAGGCAACATTAACAGCTCACCAA 170

Oy 1501 gccacaggccaggtctgcatgctctgctcccccagaggctgctgagccgagccc 1560
 db 1715 gccctgaacacacctctgcaattcttattgctctccgaagccctgagggccagcc 1774
 Oy 1561 agggagctgctctctgcccagatgctcagccagagccaggaatgctgacaagtgcaag 1620
 db 1775 agcgactgtctctctccgacgaattgtgacgagccagggagtgctggaatgcaac 1834
 Oy 1621 ctctctgagaggtgagccaaaggagttgtgagagacctgagtgacatgagtgccacca 1880
 db 1835 atctctgagagggcgaacgaagagttgtggaatcttgaattctgacatgacatcca 1894
 Oy 1681 gaggctgctctcagagccatgacacacacacacagagagagagagagagagagagag 1740
 db 1895 gaattgtctctccagccatgacacacacacacacacacacacacacacacacacacac 1954
 Oy 1741 caagtgtgcccac 1800
 db 1955 cagtggtgcccac 2014
 Oy 1801 gggag 1860
 db 2015 ggaag 2074
 Oy 1861 catcaaacctgac 1908
 db 2075 cagccac 2122

RESULT 13

HSU48722 1593 bp mRNA PRI 05-NOV-1996
 LOCUS Human epidermal growth factor receptor precursor (EGFR) mRNA,
 DEFINITION complete cds.
 VERSION U48722.1 GI:1628549
 KEYWORDS human.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 1593)
 Reiter,J.L. and Mailhe,N.J.
 A 1.8 kb alternative transcript from the human epidermal growth
 factor receptor gene encodes a truncated form of the receptor
 Nucleic Acids Res. 24 (20), 4050-4056 (1996)

JOURNAL

MEDLINE 97078686
 TITLE Reiter,J.L.
 JOURNAL Direct Submission
 Submitted (07-FEB-1996) Jill L. Reiter, Biochemistry and Molecular
 Biology, Mayo Clinic, 200 First St. SW, Rochester, MN 55905, USA
 Location/Qualifiers
 1..1593
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="p11.2-p12"
 /tissue="type-placenta"
 1..1456
 /gene="EGFR"
 1..244
 /gene="EGFR"
 245..1462
 /note="ErbB-1"
 /product="epidermal growth factor receptor precursor"
 /protein_id="A450802.1"
 /db_xref="GI:1628550"
 /translacion="KRPSTGAGTAAATLALALCAPSALAEKKVCCGTSNKLITQICTP
 EDHFLSLQRFNENCEVVLGNLITVGRNVLSTKLTQVAGVYVLAIVTERIPLE
 NLOITIRGNMYENSVYALAVLSNYDANKTGLKELPMRLQELHGAIVRFSNNPDLNVE

FEATURES

source
 gene
 5' UTR
 CDS

STGMDIVSDPLSNKMSMDPQNHUGSCQKDPSCPNKSGCAGAEENCKLTKITICAOQ
 CSCRCRKSFDSCCHNOCAAGCTGREDLVCRRKRDATCKDTCPLMLYNPTTQA
 MDVNPESKYSFEGATCYKCKPKNYVTHGSCVACACADSYEMEDGVRCKKCGPGR
 KVCNIGIIGFIFKDSLSINATNLIKHKMCTSSIDILHLPVRKDSFTHTPDLPOEL
 DILKTVKEITGLS"
 245..316
 /gene="EGFR"
 317..1459
 /product="epidermal growth factor receptor"
 626..634
 /gene="EGFR"
 /note="encodes Asn-Linked glycosylation site"
 767..775
 /gene="EGFR"
 /note="encodes Asn-Linked glycosylation site"
 830..838
 /gene="EGFR"
 /note="encodes Asn-Linked glycosylation site"
 1055..1063
 /gene="EGFR"
 /note="encodes Asn-Linked glycosylation site"
 1298..1306
 /gene="EGFR"
 /note="encodes Asn-Linked glycosylation site"
 1325..1333
 /gene="EGFR"
 /note="encodes Asn-Linked glycosylation site"
 1452..1456
 /gene="EGFR"
 /note="5' splice donor site"
 1463..1469
 /gene="EGFR"
 /note="5' splice donor site"
 1537..1542
 /gene="EGFR"
 /note="5' splice donor site"
 1569
 /gene="EGFR"
 /note="5' splice donor site"

Query Match

Best Local Similarity 59.4%; Score 1207.8; DB 9; Length 1593;
 Matches 1209; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 atggagacccctcggag 60
 db 245 atggagacccctcggag 304
 Oy 61 gggagctgagctctgag 120
 db 305 gggagctgagctctgag 364
 Oy 121 ttggagacccctcggag 180
 db 365 ttggagacccctcggag 424
 Oy 181 gggagacccctcggag 240
 db 425 gggagacccctcggag 484
 Oy 241 accagacccctcggag 300
 db 485 accagacccctcggag 544
 Oy 301 ttggagacccctcggag 360
 db 545 ttggagacccctcggag 604
 Oy 361 gggagacccctcggag 420
 db 605 gggagacccctcggag 664
 Oy 421 cagagacccctcggag 480
 db 665 cagagacccctcggag 724

QY 481 agatccagtcgagcagacatagtcagcaagtcgtctctcaagcaacatgctcagatgac 540
 |||||||
 Db 725 AGCATCCAGTGGGGGACATAGTACGACACTGACTTCTTCCACCAACATGTCATGAGACTTC 784
 QY 541 cagaacaccctggcagcctgcagcaaaagtgtatcccaagctgtcccaatggagagctgtg 600
 |||||||
 Db 785 CAGAACCACTTGCGAGCTGCCAAAAGTGTATCCAACTCTCCCATGGAGAGCTGCTGG 844
 QY 601 ggtcagagagagagagacatgcagaaactgcacaaatcatctgtgtccagcagctgtcc 660
 |||||||
 Db 845 GGTCCAGAGAGAGAGAGACATGCCAGAAACTGACCAAAATCATCTGTGCCACGACTGCTCC 904
 QY 661 gggcgcctgcctggcagagtcgcccaagtcagctgtctcacaacacagctgtgtcagctgc 720
 |||||||
 Db 905 GGGCGCTGCCGTGGCAAGTCCCGACATGACGTCTGCCACAAACAGTGTGTCGACGCTGC 964
 QY 721 acagagcccgaggagagagagagctgtgtctgtccgcaaatccagagcagagcagctgc 780
 |||||||
 Db 965 ACAGGGCCCCGGGAGAGACAGCTGCTGGTGTGCGCAAAATTCGAGAGAGAGAGAGAGAGAG 1024
 QY 781 aagagacacccctgcacccactcactcactcactcactcactcactcactcactcactcact 840
 |||||||
 Db 1025 AAGGACACCTGCCCCCCTCCTCATGCTCTACAAACCCACACAGTACAGATGATGATGATG 1084
 QY 841 cccagagagcaatacagctgtgtgtccacactgcgtgaagaagtgtcccgtaatatgtg 900
 |||||||
 Db 1085 CCCAGAGGCAAAATACAGCTTTGGTGTGCGACCTGCTGAGAAAGTGTCCCGTAATTTATGTG 1144
 QY 901 gtgcagatcaagctgt 960
 |||||||
 Db 1145 GTGACAAATACAGGCTGT 1204
 QY 961 gacgagctgcagcagctgt 1020
 |||||||
 Db 1205 GACGCGCTCCGCAAGTGAAGAGTGCAGAGGGCTTGTGCCCAAGGTGTGAACGAGATA 1264
 QY 1021 ggtattgtgtatatttaagaactcactcactcactcactcactcactcactcactcact 1080
 |||||||
 Db 1265 GGTATGTGTATTAAGACTACTCTCTCAATAAATGCTACGAATTAATAACACTTTAAA 1324
 QY 1081 aactgcacccctcactcactcactcactcactcactcactcactcactcactcactcact 1140
 |||||||
 Db 1325 AACTGCACCTTCATAGTGGCATCTCCACATCTCTGCGGTGAGATTTAGGGGTGACTCC 1384
 QY 1141 ttcaacatactctctctcctgcagatccagagagagagagagagagagagagagagagag 1200
 |||||||
 Db 1385 TTCACATATCT 1444
 QY 1201 atcaacaggtt 1211
 |||||||
 Db 1445 ATCACAGGTTT 1455

RESULT 14
 180039 180039 1868 bp DNA PAT 10-JUN-1998
 LOCATION Sequence 1 from Patent US 5708156.
 VERSION 180039
 ACCESSION 180039.1 GI:3208329
 KEYWORDS
 SOURCE
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1868)
 AUTHORS 11exis,V.V.
 TITLE Epidemal growth factor receptor-like gene product and its uses
 JOURNAL Patent: US 5708156-A 1 13-JAN-1998;
 FEATURES
 1. 1868
 location/Qualifiers
 source
 BASE COUNT 449 a 554 c 484 g 371 t
 ORIGIN

Query Match 59.4%; Score 1207.8; DB 6; Length 1868;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1209; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 atgagacccctcagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 60
 |||||||
 Db 504 ATGCGACCTCTCCGGAGCGCCGGGAGGCTCTGAGGCTGTGCTGTGCTGTGCTGTGCTGTG 563
 QY 61 gcgagctgcctcag 120
 |||||||
 Db 564 GCGAGTGGGCTCTGAG 623
 QY 121 ttgagacatttgaagatcatttctcagccctcagagagagagagagagagagagagagag 180
 |||||||
 Db 624 TTGGGCACTTTTAAATATCATTTTCTCAACCTCCAGAGAGATGTTTCAATATCTGAGGTG 683
 QY 181 gtcccttgagaaatttgaaatttacctatgtcagagagaaattgaatttcttctcttaag 240
 |||||||
 Db 684 GTCCCTGGAAATTTGAAATTAATCAATGTCAGAGAAATTAATTAATTAATTAATTAATTA 743
 QY 241 accatccagagagtgctgtgtatgttccctcactcactcactcactcactcactcactcact 300
 |||||||
 Db 744 ACATCCAGAGGTGGCTGTGTATGCTCATGTCCTCAACACAGTGAAGCAATTCCT 803
 QY 301 ttggaacacttgagatcactcagagagaaatgtactcagaaatttctatgtcttaga 360
 |||||||
 Db 804 TTGGAATACCTGCAATCAATCAAGAGAAATTAATTAATTAATTAATTAATTAATTAATTA 863
 QY 361 gtcttactaactatgtcaataataacccagagagagagagagagagagagagagagagag 420
 |||||||
 Db 864 GTCTTATCTAATCATATGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 923
 QY 421 caggaatactctgtagcgcgt 480
 |||||||
 Db 924 CAGGAATCTCTGATGCGCCGCTGCGGCTTCAAGAACACCTTGCCTGTGCAACCTGTGAG 983
 QY 481 agcatccagtgagcagacatagtcagcagtgacttctcagcaacatgtcagatgac 540
 |||||||
 Db 984 ACATCCAGTGGCGGAGACATACACACTGACTTTCACACCAACATGTCATGATGACTTTC 1043
 QY 541 cagaacaccctggcagcctgcagcaaaagtgtatcccaagctgtcccaatggagagctgtg 600
 |||||||
 Db 1044 CAGAACCACTTGCGAGCTGCCAAAAGTGTATCCAACTTCCCAATGGAGAGCTGCTGG 1103
 QY 601 ggtcag 660
 |||||||
 Db 1104 GTTCAG 1163
 QY 661 ggcgcctgcctgcagagtcgcccaagtcagctgtcacaacacagctgtgtcagagctgc 720
 |||||||
 Db 1164 GGGCGCTGGCGTGGCAAGTGTCCCATGATGCTGTCTCCCAACAGTGTGCTGCAAGGCTGC 1223
 QY 721 acagagcccgaggag 780
 |||||||
 Db 1224 ACAGGCCCCCGGAG 1283
 QY 781 aagagacacccctgcacccactcactcactcactcactcactcactcactcactcactcact 840
 |||||||
 Db 1284 AAGGACACCTGCCCCCTCCTCATGCTCTACAAACCCACACAGTACAGATGAGATGAAC 1343
 QY 841 cccagagagcaatacagctgt 900
 |||||||
 Db 1344 CCCAGAGGCAAAATACAGCTTTGGTGTGCGACCTCTGTAAGAAAGTGTCCCGTAATTAATG 1403
 QY 901 gtgcagatcaagctgt 960
 |||||||
 Db 1404 GTGACAAATACAGGCTGT 1463
 QY 961 gacgagctgcag 1020
 |||||||
 Db 1464 GACGCGTCCGCAAGTGTGAAGAGTGCAGAGGGCTTGTGCCCAAAAGTGTATACAGAGATA 1523

Db 1404 GTGACAGATCACGGCTCGTGGCTCCGAGCCGTGTGGGCCGACAGCTATGAGATGAGGAA 1463
QY 961 gacggcgtccgaagtgtaagaagtgcgaaggcccttgccgcaaatgltgaaggaata 1020
|||||
Db 1464 GACGGGCTCCGCAAGTGAAGTGGCAAGGCCCTTGCCGCAAAAGTGTGTATACGGATA 1523
QY 1021 ggtattggtgaatttaagaactcacttcctcataatgctacgaalatataacacactcaaa 1080
|||||
Db 1524 GGTATTGGTGAATTAAAGACTCACTCTCCATTAATGCTAGGAATATTAACACTTCAAA 1583
QY 1081 aactgacctccatcaatgagtcgaltccacatccctgcggtgycatttaagggtgactcc 1140
|||||
Db 1584 AACTGCACCTCCATGAGTGGCATCTCCACATCCTGCCGGGCAATTAGGGGTGACTCC 1643
QY 1141 ttacacatactcctcctctgatatccacagaactgatatlcgaaaaccgtaaggaa 1200
|||||
Db 1644 TTCACATATCTGCCCTCTGGATCCACAGAACTGGATATCTGAAACCGTAAGGAA 1703
QY 1201 atcacagggtc 1211
|||||
1704 ATCACAGGTTT 1714

Search completed: January 14, 2002, 20:13:47
Job time: 4172 sec

QY 121 ttggacatttgaagatcattctcagctccagagatgttcaatactgtgagtg 180
 Db 307 ttggacatttgaagatcattctcagctccagagatgttcaatactgtgagtg 366
 QY 181 gtccctgggaatttgaatacttaccatctgacagaggaattatgacttctcttaag 240
 Db 367 gtccctgggaatttgaatacttaccatctgacagaggaattatgacttctcttaag 426
 QY 241 accatccagaggtgctgtgtatctatctcattccataacagtgagcgaattctc 300
 Db 427 accatccagaggtgctgtgtatctatctcattccataacagtgagcgaattctc 486
 QY 301 ttggaacacctgcagatccatccagaggaataatgtactcgaataattctatgtctgca 360
 Db 487 ttggaacacctgcagatccatccagaggaataatgtactcgaataattctatgtctgca 546
 QY 361 gtcttactcaactgactgcaataataaaccggaactgaagagctgtccatgagaattta 420
 Db 547 gtcttactcaactgactgcaataataaaccggaactgaagagctgtccatgagaattta 606
 QY 421 caggaatactctgacgtgcgtgtgtgtatcagacaacacctgcgtgtgacagtgag 480
 Db 607 caggaatactctgacgtgcgtgtgtgtatcagacaacacctgcgtgtgacagtgag 666
 QY 481 agcatccaggtgcgaggaataatgacagagtgacttctcagacaatgtgcagtgacttc 540
 Db 667 agcatccaggtgcgaggaataatgacagagtgacttctcagacaatgtgcagtgacttc 726
 QY 541 cagacacacacctgcgacgtgcacaaagtgtgactccagctgtccatgagagctgtcgtg 600
 Db 727 cagacacacacctgcgacgtgcacaaagtgtgactccagctgtccatgagagctgtcgtg 786
 QY 601 ggtgtcgaagagagagagactgtccagagaataatgacaaatcatctgtgcacagtgagctcc 660
 Db 787 ggtgtcgaagagagagagactgtccagagaataatgacaaatcatctgtgcacagtgagctcc 846
 QY 661 ggtgcgtctgcgtgtgcagagtgtcccccagtgactgtgtgcacaaacagtggtgcaggtcgtc 720
 Db 847 ggtgcgtctgcgtgtgcagagtgtcccccagtgactgtgtgcacaaacagtggtgcaggtcgtc 906
 QY 721 acagagcccccgagagagagactgtcgtgtctgcacaaatctccagaaacagaaacagctgtc 780
 Db 907 acagagcccccgagagagagactgtcgtgtctgcacaaatctccagaaacagaaacagctgtc 966
 QY 781 aaggaacacctgtgcac 840
 Db 967 aaggaacacctgtgcac 1026
 QY 841 cccgagaggaataacagcttgcgtgcac 900
 Db 1027 cccgagaggaataacagcttgcgtgcac 1086
 QY 901 gtcgacaataacagctgtgcgtgtccgagacacacacacacacacacacacacacacacacacacac 960
 Db 1087 gtcgacaataacagctgtgcgtgtccgagacacacacacacacacacacacacacacacacacacac 1146
 QY 961 gacgaggtcccgagaggtgataaagatgacgaaggtcgtgcacacacacacacacacacacacacac 1020
 Db 1147 gacgaggtcccgagaggtgataaagatgacgaaggtcgtgcacacacacacacacacacacacacac 1206
 QY 1021 ggtatgtgtgaatttaagaactcactcactcaataatgtactgataataacacttcaaa 1080
 Db 1207 ggtatgtgtgaatttaagaactcactcactcaataatgtactgataataacacttcaaa 1266
 QY 1081 aactgcacacctcactcactgacatccac 1140
 Db 1267 aactgcacacctcactcactgacatccac 1326
 QY 1141 ttccacatactcactcactgacatccac 1200
 Db 1327 ttccacatactcactcactgacatccac 1386
 QY 1201 atccagaggttctgtcgtactcaggtcgtgcgtgaaacacagagacacacacacacacacacacacac 1260

Db 1387 atccagaggttctgtcgtactcaggtcgtgcgtgaaacacagagacacacacacacacacacacacac 1446
 QY 1261 gagaacacctcactcactgacatccac 1320
 Db 1447 gagaacacctcactcactgacatccac 1506
 QY 1321 gtcacacctgaacataac 1380
 Db 1507 gtcacacctgaacataac 1566
 QY 1381 gtcacacctgaacataac 1440
 Db 1567 gtcacacctgaacataac 1626
 QY 1441 ttctggacacctcctgcgtgcagaaac 1500
 Db 1627 ttctggacacctcctgcgtgcagaaac 1686
 QY 1501 gtcacacctgaacataac 1560
 Db 1687 gtcacacctgaacataac 1746
 QY 1561 aggtgcacacctcctgcgtgcagaaac 1620
 Db 1747 aggtgcacacctcctgcgtgcagaaac 1806
 QY 1621 ctctctggaggttgagacagaggttctgtgagaaacctcgtgacacacacacacacacacacacacacac 1680
 Db 1807 ctctctggaggttgagacagaggttctgtgagaaacctcgtgacacacacacacacacacacacacacac 1866
 QY 1681 ggtgtgcacctcctgcgtgcagaaac 1740
 Db 1867 ggtgtgcacctcctgcgtgcagaaac 1926
 QY 1741 cagtgtgcacctcctgcgtgcagaaac 1800
 Db 1927 cagtgtgcacctcctgcgtgcagaaac 1986
 QY 1801 gtagaataac 1860
 Db 1987 gtagaataac 2046
 QY 1861 catcacaactgtgac 1920
 Db 2047 catcacaactgtgac 2106
 QY 1921 cctcaagatcccgatcactgcagac 1980
 Db 2107 cctcaagatcccgatcactgcagac 2166
 QY 1981 gccctggagatccgctctcactgcagaaagcgcac 2034
 Db 2167 gccctggagatccgctctcactgcagaaagcgcac 2220

RESULT 4
 ID ANQ43812 standard; DNA: 6274 BP.

AC ANQ43812;
 DT 20-OCT-1993 (First entry)
 DE pRLD1D2D3; Apal construct.
 KW Epidermal growth factor receptor truncate protein; EGF; growth;
 binding sites; adsorptive agents; mammalian cell growth abnormality;
 detection; reproduction; signal transmission; pRLD1D2D3Apal; ds-
 OS Synthetic.
 FH Key Location/Qualifiers

| | |
|----------|---|
| RESULT | 6 |
| ID | AAV06408 |
| AAV06408 | standard; DNA; 1868 BP. |
| XX | |
| AC | AAV06408; |
| DT | 05-MAY-1998 (first entry) |
| DE | Epidermal growth factor receptor-like protein (TEGFR) encoding cDNA. |
| KX | Epidermal growth factor receptor-like protein; TEGFR: human; antibody |
| KX | diagnosis: cancer; hyperproliferative disease; ss. |
| US | Homo sapiens. |
| XX | |
| FH | Key |
| CDS | location/Qualifiers 504..1721 |
| FT | /*tag= a |
| FT | /product= TEGFR |
| FT | 504..575 |
| FT | /*tag= b |
| FT | /note= "putative signal peptide" |
| FT | 576..1718 |
| FT | /*tag= c |
| XX | |
| PN | US5708156-A. |
| PX | |
| PD | 13-JAN-1998. |
| PX | |
| PF | 31-MAY-1996; 96US-0658883. |
| PX | |
| PR | 31-MAY-1996; 96US-0658883. |
| PX | |
| PA | (ILEK/) ILEKIS J V. |
| PI | IleKis JV; |

This cDNA encodes a novel human epidermal growth factor receptor-like protein termed TEGFR. TEGFR is derived from human placenta. Fragments of the TEGFR cDNA can be used as a probe in a hybridisation assay to detect full length of TEGFR. An antibody can be produced using the antigen comprising at least the unique C-terminal sequence of the TEGFR polypeptide, optionally conjugated to an immunogenic carrier. The antibody can be used in an immunoassay to detect the polypeptide. The hybridisation assay and immunoassay can be used in the diagnosis of human cancers. Antibodies and other inhibitors of the TEGFR polypeptide may also be used to treat hyperproliferative diseases including cancer.

| | | | | |
|-----------------------|--------------|---------------|---------------|--------------|
| Query Match | 59.4% | Score 1207.8; | DB 19; | Length 1868; |
| Best Local Similarity | 99.8%; | Pred. No. 0; | | |
| Matches 1209; | Conservative | 0; | Mismatches 2; | Indels 0; |
| | | | Gaps | 0 |

[illegible]

Db 767 acgtcttgcctgcgtcaattcaatgacgtgagcagctgtctctcaagtgtcccaaa 826
 QY 800 lcatgtctcaaacaccacacaglaacagatgagtgaaacccagaggaacaatagcgt 859
 Db 827 ccttgcctcaaacaccacacacacacacacacacacacacacacacacacacac 886
 QY 860 ttgtgtcaccacgtgcgtgagaaagtgcccccgaattatgtgtgagaaatcagcgtcgt 919
 Db 887 atgagacatctcgtgtcagaaatgtccacataacttggc---agattccagctctc 943
 QY 920 gcgtccagaaacgtgtgagccacagcctatgagatgagaaacagcgtccgcaggtgta 979
 Db 944 gt 1003
 QY 980 agaaagtgcgaagagccttgcgtcgaagaagtgtgtaacggaataggtatgtgtgaaatc 1039
 Db 1004 aaccttgcacatgacatctgtcccaaaagcctgtgagtgagccttgacagagatcattgtg 1063
 QY 1040 aactcactcccaataaagtgtcagatatttaaacacttcaaaacgtgcacacacagtcag 1099
 Db 1064 cagctcagacgtgtggtatctccagtaacatctgaacaacttcaaacgtgtccagaatcag 1123
 QY 1100 gcgattccacatctcctcaggtgtgcatcttaaggtgtaactcctccacatctcctc 1159
 Db 1124 ggaattgtatctctctatctatctgtatctcaagtgagagacacacacacacacac 1183
 QY 1160 tgaatccacagagaaatgagatattctgaaacacgaaagaaatcacaagaggttctgtga 1219
 Db 1184 tagaaccagagaaatgagaaagcttctcgaacagtcacagagatcaagatctctcgaaca 1243
 QY 1220 ttcagccttgcgttcgaacacagagacacacacacacacacacacacacacacacac 1279
 Db 1244 taagatcagtcagcacaac 1303
 QY 1280 gcgagacagac 1339
 Db 1304 gtgtgagagatctctatctatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1363
 QY 1340 ccttgcgttgaatcagcctcctcacaagagagatgagtgagagtgagatctcagaaaca 1399
 Db 1364 ctctacagatctcagcctcctcgaagagaaatcagcagaaacatctatctcagaaaca 1423
 QY 1400 aaaaattgtgcatactcaatacaatacaatacaatacaatacaatacaatacaataca 1459
 Db 1424 gcaacctgtgtatatacaatacaatacaatacaatacaatacaatacaatacaataca 1483
 QY 1460 aaaaacaaatctataagacacagagtgaaacacagcgtcgaagacacagcagtcgtcc 1519
 Db 1484 gaatagttaattccggtgacacacagaaagctgaaatgtgtgtgtgtgtgtgtgtgtgt 1543
 QY 1520 atgacctgtgtcccccagagcgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1579
 Db 1544 acaactgtgtctcagtgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1603
 QY 1580 ggaatgtcagcagagcag 1639
 Db 1604 ggcgcctcagtgag 1663
 QY 1640 gggagattgtgagagaaatctgagtgacatagtgacacacacacagtg---ccgtgc 1696
 Db 1664 gggagattgtgagagaaatctgacatctcgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1723
 QY 1697 ccaatgacacatcagcgtcag 1756
 Db 1724 atgacctcctac 1783
 QY 1757 ttgagagccccaatctggtctcaagacacacacacacacacacacacacacacacacac 1816
 Db 1784 aagagagcccaatctggtctcaagagagagagagagagagagagagagagagagagag 1843
 QY 1817 tggatcag 1876
 Db 1844 ttctc---agatagctgtatccagatcggagagtgccacacacacacacacacacacac 1900

QY 1877 acgagatgcactgggcccaggtcttgaagcgtg 1907
 Db 1901 aaggggtgaagagctccacatgacatgactg 1931

RESULT 10
 AA064896
 ID AA064896 standard; cDNA; 5501 BP.
 XX

AC AA064896;
 XX
 DT 11-JAN-1995 (first entry)
 XX
 DE HER4.
 XX

XX Human epidermal growth factor receptor; HER; EGF; tyrosine kinase;
 KM cancer; neuronal tissue; muscle tissue; neoplasm;
 KW carcinoma; primer; probe; PCR; ss.
 XX

OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT CDS 34..3960
 FT /*tag= a

XX EP599274-A.
 XX
 PD 01-JUN-1994.
 XX

XX PE 23-NOV-1993; 93EP-0118837.
 XX
 PR 24-NOV-1992; 92US-0981165.
 XX

XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Culoussou J, Plowman GD, Shoyab M;
 XX

DR WPI: 1994-169599/21.
 DR P-PSDB: AAR54841.
 XX

PT New recombinant nucleic acid expressing HER4 - a new receptor
 PT Tyrosine kinase expressed in some cancer cells, and related
 PT vectors, antibodies, ligands etc, for diagnosis and treatment of
 PT cancers
 XX

PS Claim 4; Fig 1; 104pp: English.

XX
 CC HER4 is the fourth member of the EGFR-family of tyrosine kinases and
 CC is expressed in some human cancers and in some tissues of neuronal
 CC or muscle origin. HER4 polynucleotides, opt. labelled, are useful
 CC in assays (e.g. of HER4 mRNA to detect certain neoplasms, esp.
 CC breast carcinoma) and as primers in PCR or as probes.
 XX

SO Sequence 5501 BP; 1605 A; 1173 C; 1230 G; 1493 T; 0 other;

Query Match 21.2%; Score 431; DB 15; Length 5501;

Best Local Similarity 54.1%; Pred. No. 5.9e-113; Matches 991; Conservative 0; Mismatches 825; Indels 15; Gaps 5;

QY 83 aagaagcttgcgaagcagaglaacaaagctcagcaggttggcactttgaaagatcatt 142

Db 110 agtcagtggtgcag 169

QY 143 ttctcagcctccagagagagatctcaataactgagagtggtctcttggaatttgaaata 202

Db 170 accgagccttgcgaagtaactgaataactggaaggtgtcagtggaacacagagataa 229

QY 203 cctatgtcagagagatataatgacttctctcttlaagacatccagagagtggtcgtgt 262

Db 230 ccagatgagacacacacagagacatctctctctctcgtcgtcgtcgtcgtcgtcgtcgt 289

| Query Match | Similarity | 21.28 | Score 431 | DB 17 | Length 5501 |
|-------------|------------|---|--------------------|------------|-------------|
| Best Local | Similarity | 54.1% | Pred. No. 5.9e-113 | | |
| Matches | 991 | Conservative | 0 | Mismatches | 825 |
| | | | | Indels | 15 |
| | | | | Gaps | 5 |
| Qy | 83 | agaaagatctgcacaaagcgacgaagtaaaagctacagcagcttggcgacattttgaagatcatt | 142 | | |
| Dh | 110 | agtcacatctgctgacgaagaaacggagaaataaaacctgagctctctctctctcgaacctggaaacggcgtc | 169 | | |
| Qy | 143 | ttctcagccctccacaggaagtgctcaataactgtgagcttgagcttcttgaggaatttggaaattaa | 202 | | |
| Dh | 170 | accgagacctctgctgacaaagtactatgaaataaacttggaaagtcttgagacttgcgaacacctggagataaa | 229 | | |
| Qy | 203 | cccaatctgacaaaggaataataatgactctctctctctctaaagacacccacacggaggggtgtgtgtc | 262 | | |
| Dh | 230 | ccgagcatctgagacaacacggagacctctctctctctcgtgtcgtgtctcgagaaatcacaagcct | 289 | | |
| Qy | 263 | atctccatctgtccctccacaacacacggtgagcgagatccctctctggaaacacccgagatcatca | 322 | | |
| Dh | 290 | acgtgtctgagctgctctctatcagcttctcgttactacctgctctcgggaattttagcattatctc | 349 | | |
| Qy | 323 | gagagaaataatgtaaacctaaagaaatctctaaagctcttaagagcttatactaaactatgatacga | 382 | | |
| Dh | 350 | gtggagcaaaaactttagagagatcgtgataagcctctggacacataattttaaaccctcgaaaag | 409 | | |
| Qy | 383 | ataaaaac-----gagactgaagagagctgtgcccacatgagaaatttcaacggaaatctgcatactggcg | 439 | | |
| Dh | 410 | atggaaabacttggagacttcaagaacacttggaaattaaagaacacttgaagaataatctcaaaattggtg | 469 | | |
| Qy | 440 | ccgtctgctgtccagacaaacacccctgcccgtgacaaagtgtgagagcatctcagctgtgctggagaca | 499 | | |
| Dh | 470 | gagctcaatctgagaccaaagaaacaaatctctctgttctatggagacaacatctcatcttgagacagata | 529 | | |
| Qy | 500 | taatctcagacagatgaaacttctctacagaacaaatgctgaaatgagacactcccaagaaacccacctggcgacgc | 559 | | |
| Dh | 530 | ttgttttgagaaacccacatctcctccaaacttgaaactctgtgtctgtctcaaaataatgtttaattcaagat | 589 | | |
| Qy | 560 | gccaabaagtgatctcaacagctgtgcccaaaagtggagacactcttgggtgtgacagaaagagaaac | 619 | | |
| Dh | 590 | gtggagcgtgtgcataatagctcgtc---acgtggcgtgtgtctgtggggaacccaaagaaataatctc | 646 | | |
| Qy | 620 | gccaagaaactgaccaaataatcatctgtgtgccacggacagctgcacgggcctgcgcgtgtgacaagt | 679 | | |
| Dh | 647 | gccaagactcttcaacaagagacgtgtgtgtgacgaagaaataatgtgtgacgtgaaatgtgtactagaaacctc | 706 | | |
| Qy | 680 | cccccaagtgaactgctgacaaacaaatgtgtgactgtgacagcctgtgacaaagcctcccccggagagacg | 739 | | |

[illegible]

| | | |
|----|---|---|
| | IL/T | 13 |
| XX | A18533 | |
| ID | AAIT18533 | standard; DNA; 5555 BP. |
| XX | AC | |
| XX | AAIT18533; | |
| DT | 05-JUL-1996 | (first entry) |
| XX | | |
| DE | Receptor tyrosine kinase HER4 gene (alternative 3' sequence). | |
| XX | HER4/P180erbB4; HER4; receptor tyrosine kinase; | |
| KW | epidermal growth factor receptor; cancer; diagnosis; therapy; ss. | |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| FH | key | location/Qualifiers |
| FT | CDS | 34..3210 |
| FT | /*tag= | a |
| FT | sig-peptide | 34..108 |
| FT | /*tag= | b |
| FT | mat_peptide | 109..3207 |
| FT | /*tag= | c |
| FT | misc_difference | 4489 |
| FT | /*tag= | d |
| FT | /note= | "the base at position 4489 is given as y lnd" |

```

Query Match Summary      21.2% Score 431; DB 17; Length 5553;
Best Local Similarity   54.1%; Pred. No. 5.9e-113;
Matches 991; Conservative 0; Mismatches 825; Indels 15; Gaps 5

QY 83 agaaagttgtccaaagcagcagagataaacaagctcacagcgtctggtgcaatttggagaatc 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110 agcagagttgtgcagagaaacgagataaactgagctctctctctgctgcctgacagcagc 169

QY 143 ttctcagctccagagagatgttcaataactgtgaggtgtgctcttgggaatttggaaatta 202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 170 accgagccttcgcagaaatctactatgaataactgtgaggtttgtcagtggaacccgagata 229

QY 203 cctaatgttcgaagagaatctatgactctctctctttaaagaccatccagaggggtgtgtct 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 230 ccagagattgagacaaacgcggacctctctctcttccttcgtggtgtgtgtgtgaggttcaagc 289

QY 263 agtctcctaattgccccttaacacacagtgagcgaattcctcttggaaacccgcagatcatca 322
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 290 acgtgtctagtgtgcctctcaacagatcttcgttactctgcctccgcgggaatttaccatctac 349

QY 323 gaggaaatgtactactagaanaactctctatgcttcgaaagcgttatctactaactatgtgcaa 382
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 gtggagcaaaaacttataagagatcgaatatgactctgtgcataatttttaaaactacgaaag 409

QY 383 ataaaac--ggaactgaagagctgtcccatggagaaattctacagaaatcctcgtgtggtg 439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 atggaacacttgcactccaagaaactgtgattaaagaactctgacagaaatcctaaatgtgtg 469

QY 440 cgtgtggttcaagcaaaacccctgcgcgtgtgcacagtgagagatcccggtgtgtgtgagca 499
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 470 gactctatgtgacacagacaataatctctgttctgttctgtgacacactctatctgtgcagata 529

```


Tue Jan 15 08:29:43 2002

us-09-715-249-1_copy_1_2034.rng

Page 20

[illegible]

Search completed: January 14, 2002, 20:16:54
Job time: 4314 sec

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|----------|-----------------------|
| 1 | 1418.4 | 69.7 | 2456 | 12 | AK004911 | AK004911 Mus musculus |
| 2 | 1416.8 | 69.7 | 2662 | 12 | AK004883 | AK004883 Mus musculus |
| 3 | 1416.8 | 69.7 | 2936 | 12 | AK004944 | AK004944 Mus musculus |
| 4 | 757.8 | 37.3 | 779 | 10 | AU10518 | AU10518 AU10518 |
| 5 | 739.4 | 36.4 | 761 | 10 | AU140207 | AU140207 AU140207 |
| 6 | 671.6 | 33.0 | 960 | 10 | BE147584 | BE147584 B01576636 |
| 7 | 649.6 | 31.9 | 960 | 10 | BE172152 | BE172152 B01141248 |
| 8 | 620.2 | 30.5 | 1035 | 10 | BE142045 | BE142045 B01594768 |
| 9 | 614.4 | 30.2 | 736 | 10 | BE153882 | BE153882 B01141566 |
| 10 | 611.4 | 30.1 | 798 | 10 | BE151385 | BE151385 w966b11.x |
| 11 | 606.8 | 29.8 | 834 | 10 | BE140357 | BE140357 B01594666 |
| 12 | 569 | 29.0 | 605 | 10 | AI140805 | AI140805 w924d08.x |

| | | | | | | |
|----|-------|------|------|----|----------|---------------------|
| 13 | 531.8 | 28.6 | 757 | 10 | A1935353 | wp16907.x |
| 14 | 583.8 | 26.2 | 635 | 10 | BE262509 | BE262509 601152441 |
| 15 | 523.8 | 25.8 | 1031 | 11 | B1332447 | B1332447 602980849 |
| 16 | 515 | 25.3 | 762 | 11 | B1147112 | B1147112 602913269 |
| 17 | 509.4 | 25.0 | 678 | 10 | A1554557 | A1554557 t025f08.x |
| 18 | 502.2 | 24.7 | 784 | 10 | BE260457 | BE260457 601149610 |
| 19 | 453.4 | 22.6 | 989 | 11 | BG851511 | BG851511 353592.MA |
| 20 | 434.4 | 21.0 | 682 | 10 | A1855807 | A1855807 wj70b02.x |
| 21 | 427.8 | 21.0 | 568 | 11 | BE917493 | BE917493 601664204 |
| 22 | 423.6 | 20.8 | 641 | 10 | BE305331 | BE305331 601099975 |
| 23 | 411.8 | 20.2 | 1198 | 11 | BF347824 | BF347824 602022942 |
| 24 | 399 | 19.6 | 1059 | 11 | BG963833 | BG963833 602831495 |
| 25 | 395.4 | 19.4 | 645 | 10 | AM163375 | AM163375 au49d11.y |
| 26 | 382.8 | 18.8 | 415 | 10 | AA776320 | AA776320 z172a02.S |
| 27 | 384 | 17.9 | 540 | 10 | BE722865 | BE722865 191269.MA |
| 28 | 362.4 | 17.8 | 759 | 10 | AU139492 | AU139492 AU1394.92 |
| 29 | 349.2 | 17.2 | 558 | 10 | BE288425 | BE288425 601095327 |
| 30 | 343.8 | 17.1 | 740 | 11 | B1102081 | B1102081 602887822 |
| 31 | 333.2 | 16.4 | 464 | 10 | AN701271 | AN701271 z15f011.S |
| 32 | 331.4 | 16.3 | 488 | 11 | W14642 | W14642 md30d05.f11 |
| 33 | 327 | 16.1 | 646 | 10 | AM163038 | AM163038 au49e07.y |
| 34 | 324.6 | 16.0 | 481 | 10 | A1156625 | A1156625 au44f09.T |
| 35 | 324.4 | 15.9 | 446 | 11 | B1303786 | B1303786 602982461 |
| 36 | 321 | 15.8 | 475 | 11 | BF056131 | BF056131 71788905.x |
| 37 | 307.8 | 15.1 | 955 | 11 | BF533273 | BF533273 602073728 |
| 38 | 307 | 15.1 | 906 | 11 | BF232802 | BF232802 602023406 |
| 39 | 306.2 | 15.1 | 476 | 11 | BF058287 | BF058287 7K29a11.x |
| 40 | 302.4 | 14.9 | 764 | 10 | A1747544 | A1747544 u120a11.x |
| 41 | 299.4 | 14.7 | 430 | 11 | BF660321 | BF660321 ma29d03.S |
| 42 | 297.6 | 14.3 | 502 | 10 | A1263609 | A1263609 q989302.x |
| 43 | 280 | 14.3 | 644 | 10 | BE327106 | BE327106 h047b09.x |
| 44 | 287 | 14.1 | 664 | 10 | BE746257 | BE746257 601579743 |
| 45 | 287 | 14.1 | 935 | 10 | BE745330 | BE745330 601578113 |

ALIGNMENTS

| RESULT 1 | AK004911 | LOCUS DEFINITION | ACCESSION VERSION | KEYWORDS SOURCE | ORGANISM | REFERENCE AUTHORS | TITLE | JOURNAL MEDLINE PUBMED | REFERENCE AUTHORS | TITLE | JOURNAL MEDLINE PUBMED | REFERENCE AUTHORS |
|-------------|----------|---|---------------------------|--|--|--|--|------------------------------|---|--|------------------------------|---|
| | AK004911 | 2456 bp mRNA Mus musculus adult male liver cDNA, RIKEN full-length enriched library. Clone:1300005M11, full insert sequence. | AK004911.1 GI:12836452 | CAP Trapper. Mus musculus (strain:C57BL/6J) adult male liver cDNA to mRNA, clone:13b:RIKEN full-length enriched mouse cDNA library clone:1300005M11. | Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | 1 (bases 1 to 2456) Carrincci, P. and Hayashizaki, Y. | High-efficiency full-length cDNA cloning Methods in enzymology. 303, 19-44 (1999) | 99278253 10349636 | 2 (bases 1 to 2456) Carrincci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, H., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome research. 10 (10), 1617-1630 (2000) | 2049374 11042159 | 3 (bases 1 to 2456) Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carrincci, P., Kono, H., Aizawa, J., Nishi, K., Katsunai, T., Tashiro, T., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Tanemoto, K., Matsumoto, H., Sakaueuchi, S., Ikegami, T., Koshizaki, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J., |

| TITLE | PERSONNEL | PERIOD | REMARKS |
|--|-----------|--------|---------|
| Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y | | | |
| RIKEN integrated sequence analysis (RISA) system--384-format | | | |

JOURNAL
Genome research. 10 (11), 1757-1771 (2000)
DOI:10.1093/genres/10.11.1757

| | |
|-----------|---------------------|
| MEADLINE | 40530913 |
| PUBMED | 11076861 |
| REFERENCE | 4 (bases 1 to 2456) |

AUTHORS The RIKEN Genome Exp

| TITLE | JOURNAL | REFERENCE |
|--|----------------------------|----------------------|
| PANTOM Consortium. Functional annotation of a full-length mouse cDNA collection | Nature 409, 685-690 (2001) | 5 (change 1 to 3455) |

AUTHORS

Adachi, J., Mizawa, K., Akahira, S., Akimura, T., Aono, H., Araki, A.

Aikawa, I., Cairnini, P., Fukuda, S., Fukushima, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Imotsuki, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J.,

Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki,

Db 335 GCAGCTGGGGCGCTTGGAGGAAAGAAAGTCTGCCAAGGCAAGTAACAGCTCACCAG 394
 QY 121 tctgggactcttgaagatcatcttctcagcctcgaagagatgtcaataactgtgaagtg 180
 Db 395 ctggggacttttgaagacactttctgaagctcgaagatgtgacaaacactgtcaagtg 454
 QY 181 gtccttggaaatttgaaattactatgtgcagagaaatataatcttcttctttaaag 240
 Db 455 gtccttggaaatttgaaattactatgtgcagagaaatataatcttcttctttaaag 514
 QY 241 accatccagagagtgagctgtatctcctcctcctcctcctcctcctcctcctcctcct 300
 Db 515 accatccagagagtgagctgtatctcctcctcctcctcctcctcctcctcctcctcct 574
 QY 301 tgggaaaccccgagatcaatcagagaaatataatgttactaagaatcttcttctttaa 360
 Db 575 ttggagaaactgtgacatcattcagagaaatataatgttactaagaatcttcttctttaa 634
 QY 361 gctctatcactatgtatgtcaaaataaaccgagctgagagtgagcttcttctttaa 420
 Db 635 atccctgtccactatgtatgtcaaaataaaccgagctgagagtgagcttcttctttaa 694
 QY 421 caggaatactccgacatgagagctgtgcgttcaagaaacacccctgtgtgagagtgag 480
 Db 695 caggaatactccgacatgagagctgtgcgttcaagaaacacccctgtgtgagagtgag 754
 QY 481 agcatccagtgagcgagatagtcagacatgtacttctcagcaacatgtcgaatgttc 540
 Db 755 actatccagtgagcgagatagtcagacatgtacttctcagcaacatgtcgaatgttc 814
 QY 541 caggaatactccgacatgagagctgtgcgttcaagaaacacccctgtgtgagagtgag 600
 Db 815 caggaatactccgacatgagagctgtgcgttcaagaaacacccctgtgtgagagtgag 874
 QY 601 ggtcag 660
 Db 875 ggtcag 934
 QY 661 ggtcag 720
 Db 935 ggtcag 994
 QY 721 aag 780
 Db 995 aag 1054
 QY 781 aag 840
 Db 1055 aag 1114
 QY 841 cccgag 900
 Db 1115 cccgag 1174
 QY 901 gtcag 960
 Db 1175 gtcag 1234
 QY 961 gtcag 1020
 Db 1235 gtcag 1294
 QY 1021 gtcag 1080
 Db 1295 gtcag 1354
 QY 1081 aactcag 1140
 Db 1355 aactcag 1414
 QY 1141 ttcag 1200
 Db 1415 ttcag 1474

QY 1201 atcacaagaggttcttgcagatcaagctgtgcctgtcaaaacagagagagagagagag 1260
 Db 1475 atcacaagaggttcttgcagatcaagctgtgcctgtcaaaacagagagagagagagag 1534
 QY 1261 gagaacacacag 1320
 Db 1535 gagaacacacag 1394
 QY 1321 gagaacacacag 1380
 Db 1595 gagaacacacag 1654
 QY 1381 gtcag 1440
 Db 1655 gtcag 1714
 QY 1441 ttcgag 1500
 Db 1715 ttcgag 1774
 QY 1501 gtcag 1560
 Db 1775 gtcag 1834
 QY 1561 gtcag 1620
 Db 1835 gtcag 1884
 QY 1621 gtcag 1680
 Db 1895 gtcag 1954
 QY 1681 gtcag 1740
 Db 1955 gtcag 2014
 QY 1741 gtcag 1800
 Db 2015 gtcag 2074
 QY 1801 gtcag 1860
 Db 2075 gtcag 2134
 QY 1861 gtcag 1920
 Db 2135 gtcag 2182

RESULT 4
 LOCUS AU140518 779 bp mRNA EST 25-OCT-2000
 DEFINITION AU140518 PLACE3 Homo sapiens CDNA clone PLACE300215 5', mRNA
 sequence.
 ACCESSION AU140518
 VERSION AU140518.1 GI:11002039
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 779)
 AUTHORS Ota, T., Suzuki, Y., Saito, K., Ishii, S., Yamamoto, J., Sugiyama, T.,
 Nishikawa, T., Nakamura, Y., Sugano, S., Masuno, Y. and Isogai, T.
 HRI human cDNA project (Ota, T., Suzuki, Y., Saito, K., Ishii, S.,
 Yamamoto, J., Sugiyama, T., Nishikawa, T., Nakamura, Y., Sugano, S.,
 Masuno, Y., Isogai, T.)
 JOURNAL Unpublished (2000)
 COMMENT Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3951
Fax: 81-438-52-3952

Email: genomicshri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES

source
1. .779
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="PLACE3000215"
/clone_id="PLACE3"
/issue_type="placenta"
/note="Vector: pME18SFL3"
BASE COUNT 208 a 207 c 205 g 153 t 6 others
ORIGIN

Query Match 37.3%; Score 757.8; DB 10; Length 779;

Best Local Similarity 98.4%; Pred. No. 6, 4e-187; Indels 0; Gaps 0;
Matches 762; Conservative 0; Mismatches 12;

```

QY 318 catcagaggaatatactactagaataatctccttagcagctcttatactataatga 377
    |||
Db 2 CATCAGAGGAAATATGTCATCGAAATTCCTATGCTTAGCAGCTTATCTAACTATGA 61
    |||

QY 378 tgcataataaacccgactggaagagctgcccactgagaatattacaggaatctgcatg 437
    |||
Db 62 TGCATAATAAACCCGACTGGAAGAGCTGCCATGACAAATTTACAGAAATCTGTCATGG 121
    |||

QY 438 cgcgcgtggttgaagaaacccctgctgctgaacgtggaagagctatccagtgagga 497
    |||
Db 122 CGCGCGGCGTTCACACAAACCTGCGCTGTCACAGTGAGAGATCCATCCATGCGGGA 181
    |||

QY 498 catactgagcagctgactcttctcagcaaatgctgatatgacttcagaaaccttgagag 557
    |||
Db 182 CATAGTCAGCAGTGTACTTTCACCAACATGATGATGAGATTCACGAAACACCTCGGACG 241
    |||

QY 558 ctgcacaaagctgtgacaaagctgtcccaatggaagctgctgggtgacagagagaa 617
    |||
Db 242 CTGCCAAAGTGTGATCCCAACCTGTCCCATGGAGACTGCTGGGTGACAGAGAGAGAA 301
    |||

QY 618 ctgcagaagaatgacaaatcatctctgcccagcagctgctcgggcgctgcgctgagaa 677
    |||
Db 302 CTGCCAAGAAATGACCAAAATCATCTGTGCCAGAGTGTGCTCGGCGCTGCGTGGCA 361
    |||

QY 678 gtcccccagctgactgctgcacaaacagctgtgctgcaagctgcacagagcccccggagag 737
    |||
Db 362 GTCCCCAGTGTACTGCTGCCACACAGTGTGCTGCAAGGCTGCACAGGCCCCCGGAGAG 421
    |||

QY 738 gactgctgctgtctgcacaaatctcgaagacagcagctgcaagagaaacctgcccccc 797
    |||
Db 422 CGAGTGTCTGTCTGCTGCCCAATTCGAGACGAAGCCAGCTGCMAGGACCTGCCCCC 481
    |||

QY 798 actcatgctctacaaacccacacgtaacagatgagatgaaaccccgaggaataacag 857
    |||
Db 482 ACTCATGCTCTCAACCCACACAGTACAGATGATGTAACCCGAGAGCAATTAACAG 541
    |||

QY 858 ccttctgtgcaacctgctggaagagctgcccgaatlaatgtagtaacagatcaagagct 917
    |||
Db 542 CTTTGTGTCACACTGCTGCGAAGAGTGTCCCGGAATTAATGTTGTCAGATCAGGATC 601
    |||

QY 918 gtagctgcgaagctgtgagggcgaacactatgagatgagagaaagagctgcccgaatg 977
    |||
Db 602 GTAGCTGCAGAACCTGTGGGCGCGAGCTATGAGATGAGAGAGAGAGAGAGAGAGAG 661
    |||

QY 978 taagaagctgcaagagctgtgcgcaagagctgtaacagaaatgagatgagatgagaa 1037
    |||
Db 662 TAAGAAGCTGCCAAGAGCTTCCCGCAAGTGTGAACGGAATGAGATGAGATTTAA 721
    |||

QY 1038 agactactctcattaaatgctacgaatatlaaacacttcaaaactgcaacct 1091
    |||

```

DB 722 AGACTCACTMTTCATAAATGCTACNAATATTAACCCCTTTAAAAACTGGCCTTC 775

RESULT 5
LOCUS AUI140207 761 bp mRNA EST 25-OCT-2000
DEFINITION AUI140207 PLACE2 Homo sapiens cDNA clone PLACE2000123 5', mRNA
sequence.
ACCESSION AUI140207
VERSION AUI140207.1 GI:11001728
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 761)
Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T.,
Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isogai,T.,
HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,
Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,
Masuho,Y., Isogai,T.)

TITLE

JOURNAL Unpublished (2000)
CONTACT Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomicshri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES

source
1. .761
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="PLACE2000123"
/clone_id="PLACE2"
/issue_type="placenta"
/note="Vector: pME18SFL3"
BASE COUNT 209 a 199 c 199 g 150 t 4 others
ORIGIN

Query Match 36.4%; Score 739.4; DB 10; Length 761;
Best Local Similarity 98.0%; Pred. No. 4e-182; Indels 0; Gaps 0;
Matches 746; Conservative 0; Mismatches 15;

```

QY 1080 aaatgcaccccaatcagtgagcgtatccacacccctgcgcggtgagatctagggtagctc 1139
    |||
Db 1 AATCGGACCCCTCATCATGATGGGATCTCCACATCGCGCGGTGGCATTTAGGGTACTC 60
    |||

QY 1140 ctccacatctctctctctctctgataccagaagactgagatctcgaaaacaggaagga 1199
    |||
Db 61 CTTTCACATATCTCTCTCTCTGATCCACAGAGACTGGATTTCTGAAGAGAGAGAGAG 120
    |||

QY 1200 aatccagaggtttctgtgatacagccttgagcctgaaacagagagagcctcctcctt 1259
    |||
Db 121 AATCAGAGGCTGTTGCTGATTCAGGCTTGCCCTGAAACAGAGAGGACCTCCATGCTT 180
    |||

QY 1260 tgaagacctagaataatcagcagcagcagcagcagcagcagcagcagcagcagcagc 1319
    |||
Db 181 TGAGAACTAATAATCTATACCGCGCAGACAGACAGACAGATGTCAGTTCTCTTGCAAG 240
    |||

QY 1320 cgtcagcctgagaataacatctctgagatgaagctcctccctcaagagagataatgtagaga 1379
    |||
Db 241 CGTCAGCTCTGAGATTAATCATCTTGGATTAAGCTCTCAAGAGAGATTAATGATGAGAG 300
    |||

QY 1380 tctgataattccaagaacaaatattgtgctatgtaacatacaataaacttgagaaact 1439
    |||
Db 301 TGTGATTAATTTCAAGAAACAAATTTGCTGCTATGCAAAATACATTAATCTGAAAAAAT 360
    |||

```

ORIGIN

| | | | | |
|---------------------------|--------|---------------------|-----------|-------------|
| Query Match | 33.08; | Score 671.6; | DB 10; | Length 969; |
| Best Local Similarity | 99.3%; | Pred. No. 2.2e-164; | | |
| Matches 685; Conservative | 0; | Mismatches 4; | Indels 1; | Gaps 1 |

| | | | |
|------------|---|--|------------------------------|
| OY | 1345 | ggaatagctccccaaggagaatgaatgagaaatgataatctcgaataacaaat | 1404 |
| Db | 24 | GGATTAACGCTCTCTCAGAGGAGTAAGTATGATGAGATGTATATTTCAGGAACAAAT | 83 |
| OY | 1405 | tgtgtctatgcaaatataataaacctgggaataactgtttgggaactccggtcgaataacc | 1464 |
| Db | 84 | TTTGTGCTATGGAAATACATAAATCTGGAAAAACTGTTGGAGACTCCGGTCAGAAAACC | 143 |
| OY | 1465 | aaatataaagcaacaaaggatgaataacagctgtcaaggacacagggccaaggtctgcattgc | 1524 |
| Db | 144 | AAATATATTAACCAACACAGGTGCAAAACAGCTGCACAGGCCACAGGCTGTGCATGGC | 203 |
| OY | 1525 | tgtgtctcccccgagagctgcctcgagggcccgagagcccaaggactgtgtctcttcggaat | 1584 |
| Db | 204 | TTGTGCTCCCCGAGGGCTGCTGTGGGGCCCGAGGCCAGGAACTGTGCTCTTTCGCGAAT | 263 |
| OY | 1585 | gtcagccgagcgcaaggaaatgctgtggacaagtgtcaagcttcggagagtgagccaaggag | 1644 |
| Db | 264 | GTCACCCAGAGCAAGCAATGCTGTGACAAATGCAACCTTCTGAGAGGTGAGCAAGGAG | 323 |
| OY | 1645 | tttgtgtgagaaactctgagtgacaataagtcacacccagagtgctgtccttcagagcaatgaac | 1704 |
| Db | 324 | TTTGTGTGGAAGACTCTGAGTGTGATACAGTGCACCCAGACTGCTCTCCATCAGGCCATGAAC | 383 |
| OY | 1705 | atcaccttcacaagaacggggagacacgaataacgttatctcaatgtgtccacatacttgaagc | 1764 |
| Db | 384 | ATCACCCTGCACAGSAGCGGGACCAACACATCTATCTCAATGTGCCCATATTGAGCGC | 443 |
| OY | 1765 | ccccacatctcgttcaagactgtcccgagaaagtacaatggaaataaacaccctgtgtctg | 1824 |
| Db | 444 | CCCCACTCTGTCACAGACATGTCCTCCGGCAGAGATATATGGAGAAAACACCCTGGTCTTG | 503 |
| OY | 1825 | aagtaacgaagacgtgcggcgaatgtgtgcacactgtgcataccaatactgaactctagatgc | 1884 |
| Db | 504 | AAGTACCCGACAGCGCGGCATGTGTGCCACTGTGCCATCCAAACTGCACCTACGGATGC | 563 |
| OY | 1885 | actggagcgaaggtcttaagaagctgtccaaagaatgggacctaaataatccgttcactcgcaact | 1944 |
| Db | 564 | ACTGGGCCAGGTCTTAAAGGCTGTCAACGAATAGGGCTCTAAGATCCGTCAATCGGCACCT | 623 |
| OY | 1945 | gggatagtgtgggggcctctctcttctgtcgtggtgtgtgtggccctggagatcgactctcaag | 2004 |
| Db | 624 | GGATAGTGGGGGCGCCCTCCCTCTTGTGTGTGTGTGTGTGGCCCTGGGATCGG-CTTCATATG | 682 |
| OY | 2005 | cgaaagcgccacatctgttcggaagcgcaag | 2034 |
| Db | 683 | CGAAGCGCCACATCGTTGGAGAGCGCAGC | 712 |
| RESULT | 7 | | |
| LOCUS | BE272152 | 960 bp | mRNA |
| DEFINITION | 60114248F1 NIH_MGC_9 | Human sapiens | cdna clone IMAGE:3141346 5', |
| ACCESSION | BE272152 | | |
| VERSION | BE272152.1 | GI:9146475 | |
| KEYWORDS | EST. | | |
| SOURCE | human. | | |
| ORGANISM | Human sapiens | | |
| REFERENCE | | | |
| AUTHORS | Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| TITLE | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| JOURNAL | NIH-MGC http://mhc.nci.nih.gov/ | | |
| COMMENT | National Institutes of Health, Mammalian Gene Collection (MGC) | | |
| | Unpublished (1999) | | |
| | Contact: Robert Strausberg Ph.D. | | |

Email: cgaabs-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at: image.lnl.gov
 Plate: LNCMI10 row: o column: 11
 High quality sequence start: 32
 High quality sequence stop: 751.
 Location/Qualifiers

FEATURES

source

1. 960
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3141346"
 /clone_id="NIH_MGC_9"
 /tissue_type="adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: ovary; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dt priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCAAGAG(G). Size selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 209 a 256 c 306 g 189 t
 ORIGIN

Query Match

Best Local Similarity 99.1%; Score 649.6; DB 10; Length 960;
 Matches 674; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1356 cctcaagagagataagtgatgagatgataatttcaggaacaaatttgctatg 1415
 11
 Db 46 CCTCTGAGGATTAAGTGTGATGATGATTAATTTCAAGAAACAAATTTGCTATGC 105
 QY 1416 aatatcaataacatgaaactgtttgagctccgcgcagaaacaaataataag 1475
 11
 Db 106 AATATACAAATAAGTGAAGAAACCTTTGGAGCTCCGTCAGAAACCAATTTATAG 165
 QY 1476 caacaagatgaaacagtgtaagtgacagagcagagtgctgcatgcttgcctcc 1535
 11
 Db 166 CAACAGAGGAGTAAGTGTGATGATGATTAATTTCAAGAAACAAATTTGCTATGC 225
 QY 1536 cgaagctcgtgagccgagagccagagactgctccttcgaggaatgcaagccag 1595
 226 CGAGGCTGCTGGGCGCCGAGACCCAGGAGCTGCTCTTCCGCGAATGTCAGCCGAG 285
 QY 1596 caaggaatcgatgagaaagtcaagcttctgagagtgatgagcaaggaagtttgtga 1655
 286 CAGGGAATCGTGAGCAAGTGCACACCTTCTGAGAGTGACCAAGGAGTGTGTGAGCA 345
 QY 1656 cctcgtgatacagtgccaccagagtgctcctcctcagagcagtaacacactgac 1715
 346 CTCTGAGTGCATAGTGCACACCTTCTGAGAGTGACCAAGGAGTGTGTGAGCA 405
 QY 1716 agaacgaggaacgaacatgtatcagtgctgacactacatgagagcccaactgct 1775
 406 AGGACGGGAGCAGCAACTGATATCCAGTGTGACCACTACATTTAGCGCCCGAC 465
 QY 1776 caagacttcgagcagagatgataagtgagaaacaaacacactgctgttgaagtaagcaga 1835
 466 CAAGACTGCGCGCAGAGTGTATGAGGAGAAACAAACACCTTGTGAAAGTACCGACA 525
 QY 1836 cgcgcgcatgtgtccactgtgcatcacaactgacactacagatgacatgagccag 1895
 526 CGCGCGCATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 585
 QY 1896 tcttgaaggtgtcacaagatgagcctaaatccgltcactgcagccagagtgatg 1955
 586 TCTTGAAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645

QY 1956 ggcctcctctgctgctgagtgagccctgaggtcgcctcttc-atcgcaagcgcc 2014
 11
 Db 646 GCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705
 QY 2015 acatgctcgaagcagcagc 2034
 11
 Db 706 ACATCG-TCGCAAGCGCAGC 724

RESULT 8

BE742045 1035 bp mRNA EST 15-SEP-2000
 LOCUS 601594769P1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3948761 5',
 DEFINITION mRNA sequence.
 ACCESSION BE742045
 VERSION BE742045.1 GI:10156037
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 1035)
 NIH-MGC http://mgc.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at: image.lnl.gov
 Plate: LNCMI12 row: l column: 18
 High quality sequence start: 64
 High quality sequence stop: 775.
 Location/Qualifiers

FEATURES

source

1. 1035
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3948761"
 /clone_id="NIH_MGC_9"
 /tissue_type="adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: ovary; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dt priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCAAGAG(G). Size selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 248 a 264 c 320 g 202 t
 ORIGIN

Query Match

Best Local Similarity 98.0%; Score 620.2; DB 10; Length 1035;
 Matches 681; Conservative 0; Mismatches 8; Indels 6; Gaps 5;

QY 1341 cttgagatgagctccctcagaag-agataagtgatgagatgataatttcaggaaga 1399
 11
 Db 37 CTGGGATTTAGCTCCCTCCCAAGAGTAAGTATGATGATGATTAATTTCAAGAAACA 96
 QY 1400 aaaaattgtctatgcaataataacatgaaacacgttcttgagaccccgatcaga 1459
 11
 Db 97 AAAATTTGTGTATGCAATTAATTAATGAGAAACATGTTGTGACCTCCGTCAGA 156
 QY 1460 aaacacaaatataagaacagagtgagaaacagctgcagagccagagctcgcgc 1519
 11
 Db 157 AAACCAAAATTTATTAAGCAACGAGGTGTAAACAGCTGCAAGGCCACAGCGCTGCGC 216

| QY | 1520 | atgaccttgcgtctccgcagagggcgtgcgtggggcccgagacgaaggagctgcctctgc | 1579 |
|------------|---|--|------|
| Db | 217 | atgcctgttgacgtcccccgcagggcgtctggggcccgagacca-ggactgcgtctcttgc | 275 |
| QY | 1580 | ggaaattcgaagccgaaggaggaatgcgttgagacaagtgcgaagctcttgaaaggtgaagcaa | 1639 |
| Db | 276 | gcattgcacaccacagacagacgaatgccttgacacaaatgcacacttctggagagtgacgcaa | 335 |
| QY | 1640 | gggaagctctgtggaagaaacctgcagtgcaatacagtcacaccagaagtgcctgcacagcca | 1699 |
| Db | 336 | ggcagatttcttgagaaactctcagtgcatatcagtgccaccacagatgcctctcagccca | 395 |
| QY | 1700 | tgaaacataccacgcagagagagggagacagaacagctctcagtgctgcacataatgtg | 1759 |
| Db | 396 | tgaaacatcactgcacacagacaggggaccacaaactgtatccagtgctccacacttg | 455 |
| QY | 1760 | acggagcccaatgcgtctaaagaccgtgcggcgagagctcttgagagaataaacacaccttg | 1819 |
| Db | 456 | acggagcccaatgcgtctaaagaccgtgcggcgagagctcttgagagaataaacacaccttg | 515 |
| QY | 1820 | tcctggaagtcagcagagcgcgggcaatgtgtgtccacctgtgtgcatacctaactgcactacg | 1879 |
| Db | 516 | tcctggaagtcagcagacacgcggcgatgtgtgtccacctgtgtgcatacctaactgcactacg | 574 |
| QY | 1880 | gatgcacctggagcagagctcttgaaagcgtgtccaaagaaatggagctcagaatctccgcctatg | 1939 |
| Db | 575 | gatgcacctggagcagagctcttgaaagcgtgtccaaagaaatggagctcagaatctccgcctatg | 634 |
| QY | 1940 | ccactgtggaatgtgtggggccctctctctgtctgtgtgtgtgccttggaatgcagcctct | 1999 |
| Db | 635 | ccactgtggaatgtgtggggccctctctctgtgtgtgtgtgtgcct--tcgggattgcgctct | 692 |
| QY | 2000 | tcatgcgaagcgcacacatcgtctcgaaagcgcacg | 2034 |
| Db | 693 | tcattgcgaag--gcacacactcttcgaaagcgcacg | 726 |
| RESULT | 9 | | |
| LOCUS | BEJ15382 | | |
| DEFINITION | 60114156881 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3141184 5', | | |
| ACCESSION | BEJ15382 | | |
| VERSION | BEJ15382.1 | | |
| KEYWORDS | EST. | | |
| SOURCE | human. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | |
| TITLE | 1 (bases 1 to 736) | | |
| JOURNAL | NIH-MGC http://mgi.nci.nih.gov/ | | |
| COMMENT | Unpublished (1999) | | |
| | Contact: Robert Strausberg, Ph.D. | | |
| | Email: cgaaps-retail.nih.gov | | |
| | Tissue Procurement: DCFD/DPF | | |
| | cDNA Library Preparation: Ling Hong/Rubin Laboratory | | |
| | cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNU) | | |
| | DNA Sequencing by: Incyte Genomics, Inc. | | |
| | Clone distribution: MGC clone distribution information can be | | |
| | found through the I.M.A.G.E. Consortium/LNU at: image.lnl.gov | | |
| | Plate: UCM10 row: h column: 17 | | |
| | High quality sequence stop: 627. | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..736 | | |
| | /organism="Homo sapiens" | | |
| | /db_xref="taxon:9606" | | |
| | /clone="IMAGE:3141184" | | |
| | /clone_1lb="NIH_MGC_9" | | |
| | /tissue_type="adenocarcinoma cell line" | | |
| | /lab_host="DH10B (phage-resistant)" | | |
| | /note="Organ: ovary; Vector: pOT5; Site_1: XhoI; Site_2: | | |
| | EcoRI; cDNA made by oligo-dT priming. Directionally | | |

cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Size-selected >50bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).

| | | | | |
|-----------------------|--------|--------------|---------|-------------|
| Query Match | 30.2%; | Score 614.4; | DB 10; | length 736; |
| Best Local Similarity | 98.8%; | pred No 1 | 7a-149. | |

| | | | | | | | | | |
|---------|------|--------------|----|------------|----|--------|----|------|----|
| Matches | 640; | Conservative | 0; | Mismatches | 6; | Indels | 2; | Gaps | 2, |
|---------|------|--------------|----|------------|----|--------|----|------|----|

QY 1369 agtcatgagagctgtataatttcaggaacaaaaatttctgctatgcataac-aataaa 1427

DB / AGTGAATGGAGATGTGATTAATTTTCAGGAACAATAAATTTCTGCTATGCAAGAAGCAGATAAA 66

1420 ccggaataaacctcgtctcgggacccctcgggtccaggaaacccaaatctataagcaacagaggtcga 148/

[illegible][illegible][illegible]

Pb 187 GGGCCGGAGCCAGGACTGCTCTTTGCGGATGTCAGCCAGGCAATGCCCT 246

QY 1608 qqacaaqtacaaqcttctqaaqqotaaqccaaqaaatttataqaaqaacctctaaqatcat 1667

Db 247 GGACAGTGCACCTTCTGGAGGGTGA6CCAGGAGATTGTGGAGACCTGAGTGCAT 306

QY 1668 acagtgcacacccagagtgctgcctcagcgcatagaacatacactgcacacagagacgagacc 1727

Db 307 ACAGTCCACCCAGAGTGCCTCAGGCCATGAACATCACCCTGCACAGAGCGGGACC 366

QY 1728 agacaactgtatccagtggtgcccactacatctgacggcccccaactgctcaaacctgcc 1787

Db 367 AGACAACTGTATCCAGTGTGCCCACTACATTTGACGGGCCCCCACTGGGTCAAGACCTGCC 426

OY 1788 ggcagagagtcattgggagaaaaaacacccctggctctgaaagtaacgcagacgccggccattgt 1847

Db 427 GGCAGAGTCATTGGGAGAAACACACCCCTGGTCTGSAAGTACGCAGACGCCGCCCATGT 486

[illegible]

1000 240

[illegible]

On 1968 actactatnatnatnaccctnnnnatcnccctcrtcatncnaaanncnca 2015

Db 607 GCTGCTGCTGAGGCTGAGGATCGG-CTTCTTCTTTCGAGAGGCCCCA 653

RESULT 10
A1761585

| | | | | | |
|------------|------------|-----------------------------|--------------|------------|-------------|
| LOCUS | AI761363 | 798 bp | miRNA | ESI | 20-DEC-1999 |
| DEFINITION | wg66b11.x1 | Soares_NSF_F8_9W_OT_PA_P_S1 | Homo sapiens | CDNA clone | |

RECEPTOR PRECURSOR (HUMAN);, mRNA sequence.

```

VERSION          AI761585.1  GI:5177252

```

| | |
|----------|--------|
| SOURCE | human. |
| ORGANISM | human. |

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo

REFERENCE 1 (bases 1 to 798)


```

|||||
Db 61 AATTAACGTGAAAAAAGTGTTCGACCTCCGTCAGAAAACAAATTTAAGCAAG 120
Oy 1482 aggtgaaacacgctgcaagacagagcaggttctgcaatgcttctgctcccgagag 1541
Db 121 AGGGAAGAACACTCTCAAGGCGCACAGGCGAGTCTGTCATGCTTGTCTCCGAGAG 180
Oy 1542 ctgctgagggccgagagcagagctgctctctgctgagaaatgctgagcagagagga 1601
Db 181 CTCTGGGCGCCGAGAGCCAGGAGCTGCTCTTGTCCGGAATGTCAAGCCAGAGGGA 240
Oy 1602 atgctgagagagctgcaagctctgagagagtgagcagagagagtgagagagagagc 1658
Db 241 ATGCGGAGCAAGTGCACACTTCTGAGAGGTGAGCCAGGAAGTTCGTGCGAGAAACTC 300
Oy 1659 tggatgatacagctgcaagagctgctgctgctgagagcagagagcagagagag 1718
Db 301 TGAGTGCATACAGTGCACAGGAGTCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 360
Oy 1719 agcggagacagacacatgctgctgctgctgctgctgctgctgctgctgctgctgct 1778
Db 361 AGGGAAGAACACTCTCAAGGCGCACAGGCGAGTCTGTCATGCTTGTCTCCGAGAG 420
Oy 1779 gacctgagcagagagagagagagagagagagagagagagagagagagagagagag 1838
Db 421 GACCTGCGGCGAGGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Oy 1839 cggcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1898
Db 481 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Oy 1899 tgaagagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1958
Db 541 TGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Oy 1959 cccctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1992
Db 601 CCTCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 634

RESULT 12
LOCUS A1740805 805 bp mRNA EST 19-DEC-1999
DEFINITION wgzad08.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
IMAGE:2366031.3 similar to gb:K03193 EPIDERMAL GROWTH FACTOR
RECEPTOR PRECURSOR (HUMAN); mRNA sequence.
ACCESSION A1740805.1 GI:5109093
VERSION 1
GDS 1
ORIGIN
human.
Homo sapiens
Mammalia: Eutheria: Primates: Catarrhini: Homidae: Homo.
1 (bases 1 to 805)
NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaops@mail.nih.gov
This clone is available royalty-free through LNL: contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1250 SCD Error: 0.00
Seq primer: -400p from Gibco
High quality sequence sloop: 488.
Location/Qualifiers
1. 805
/Organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2366031"
/clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: p773D-pac (Pharmacia) with

```

```

a modified polylinker. Site_1: Not 1; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NBSF pool 1:
309384-310919, 323208-325895 Soares N2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 150002 -
150407, 151176-152327 Soares N2HP-9W pool 1:
758280-760583, 772104-774407 Soares N2HPA pool 1:
304776-306311, 320136-322823, 326580-326653 Soares NBSF
pool 1: 723720-726407, 739080-740999 Subtraction by Benito
Soares and M. Fatima Bonaldo.
BASE COUNT 188 a 227 c 222 g 160 t 8 others
ORIGIN
Query Match 29.0%; Score 589; DB 10; Length 805;
Best Local Similarity 95.4%; Pred. No. 7,4e-143;
Matches 623; Conservative 0; Mismatches 28; Indels 2; Gaps 2;
Oy 1 atcgaccctccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 60
Db 155 ATGCAACCTCTCCGAGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 214
Oy 61 ggcagctgagcctcggagagagagagagagagagagagagagagagagagagagagagag 120
Db 215 GCGAGTCGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 274
Oy 121 ttgagacatcttcaagatcatcttctcagcctcagagagagagagagagagagagagag 180
Db 275 TTGGGCACTTTTGAATCATCTTCTCAGCCTCAGAGAGATTTGAATCATCTTCTCAGAG 334
Oy 181 gtctctgagagatctgagagatctcagagagagagagagagagagagagagagagagag 240
Db 335 GTCTTGGGAGATTGGAATTTACATATGTCAGAGAGATTTGATCTTCTCTTTAAGG 394
Oy 241 accatccagagagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 300
Db 395 KCCATCCAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 454
Oy 301 ttgagagacctgagagagagagagagagagagagagagagagagagagagagagagag 360
Db 455 TTGGAAGACCTGAGATCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 514
Oy 361 gtctatctaatctgagagagagagagagagagagagagagagagagagagagagagagag 420
Db 515 GTCTATCTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 574
Oy 421 cagagagacctgagagagagagagagagagagagagagagagagagagagagagagag 480
Db 575 CAGGAGATCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 633
Oy 481 agcatccagagagagagagagagagagagagagagagagagagagagagagagagagag 540
Db 634 ACATATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 693
Oy 541 cagagacacctgagagagagagagagagagagagagagagagagagagagagagagag 600
Db 694 CAGAACCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 752
Oy 601 ggtcagagagagagagagagagagagagagagagagagagagagagagagagagagag 653
Db 753 GGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 805

RESULT 13
LOCUS A1935313 757 bp mRNA EST 08-MAR-2000
DEFINITION w016607.x1 NCI-CCAP.Lu19 Homo sapiens cDNA clone IMAGE:2465052.3'
similar to gb:K03193 EPIDERMAL GROWTH FACTOR RECEPTOR PRECURSOR

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2002, 20:13:47 ; Search time 2222.52 Seconds
(without alignments)
534.437 Million cell updates/sec

Title: US-09-715-249-1_COPY_1_72
Perfect score: 72
Sequence: 1 atgcgacctccggagcgc.....ctgcgccggcgaatcgagct 72

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

- 1: gb_ha:*
- 2: gb_hlg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_om:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_pat:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_sy:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htgo_hum:*
- 31: em_htgo_inv:*
- 32: em_htgo_rnd:*
- 33: em_htg_hum:*
- 34: em_htg_inv:*
- 35: em_htg_rnd:*
- 36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

8

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|------------|--------------------|
| 1 | 72 | 100.0 | 560 | 9 | HUMEGFRG | M11234 Human epide |
| 2 | 72 | 100.0 | 1347 | 9 | HSEGRF1 | X06370 Human A431 |
| 3 | 72 | 100.0 | 1593 | 9 | HSU48722 | U48722 Human epide |
| 4 | 72 | 100.0 | 1868 | 6 | I80039 | I80039 Sequence 1 |
| 5 | 72 | 100.0 | 1868 | 9 | HSU95089 | U95089 Human trunc |
| 6 | 72 | 100.0 | 2643 | 9 | HUMEGFRS | K03193 Human abert |
| 7 | 72 | 100.0 | 2864 | 9 | AF125253 | AF125253 Homo sapi |
| 8 | 72 | 100.0 | 3633 | 6 | AX179384 | AX179384 Sequence |
| 9 | 72 | 100.0 | 3888 | 9 | HUMEGFR | M38425 Human EGF r |
| 10 | 72 | 100.0 | 3888 | 11 | G31681 | G31681 SWSS1476 tr |
| 11 | 72 | 100.0 | 5532 | 6 | AR086089 | AR086089 Sequence |
| 12 | 72 | 100.0 | 5532 | 9 | HSEGRFE | X00588 Human mRNA |
| 13 | 72 | 100.0 | 177998 | 9 | AC006977 | AC006977 Homo sapi |
| 14 | 72 | 100.0 | 197496 | 9 | AF288738 | AF288738 Homo sapi |
| 15 | 43.8 | 60.8 | 1858 | 10 | AB025197 | AB025197 Rattus no |
| 16 | 43.8 | 60.8 | 1958 | 10 | AF187818 | AF187818 Rattus no |
| 17 | 43.8 | 60.8 | 3208 | 10 | AF142153 | AF142153 Rattus no |
| 18 | 43.8 | 60.8 | 4194 | 10 | RATEGER | M37394 Rattus norv |
| 19 | 42.2 | 58.6 | 1292 | 10 | AF275364S1 | AF275364 Mus muscu |
| 20 | 42.2 | 58.6 | 2301 | 10 | MMBGR | X59698 M.musculus |
| 21 | 42.2 | 58.6 | 2618 | 10 | AF124513 | AF124513 Mus muscu |
| 22 | 42.2 | 58.6 | 5935 | 10 | AF275367 | AF275367 Mus muscu |
| 23 | 40.6 | 56.4 | 3704 | 10 | MMU03425 | U03425 Mus musculu |
| 24 | 40.6 | 56.4 | 4188 | 10 | MMBGR | X78987 M.musculus |
| 25 | 40.6 | 56.4 | 163451 | 2 | AC069468 | AC069468 Mus muscu |
| 26 | 33.6 | 46.7 | 1775 | 9 | AF115392 | AF115392 Homo sapi |
| 27 | 33.6 | 46.7 | 19527 | 2 | AC067859 | AC067859 Homo sapi |
| 28 | 32.6 | 45.3 | 1280 | 6 | ARI28142 | ARI28142 Sequence |
| 29 | 32.6 | 45.3 | 2648 | 5 | GCHSP905 | X15028 Chicken hsp |
| 30 | 32.6 | 45.3 | 16188 | 1 | AE006968 | AE006968 Mycobacte |
| 31 | 32.6 | 45.3 | 28826 | 1 | MTV041 | AL021958 Mycobacte |
| 32 | 31.8 | 44.2 | 2223 | 1 | MXA133131 | AL133131 Myxococcu |
| 33 | 31.4 | 43.6 | 35654 | 1 | SC7H1 | AL021411 Streptomy |
| 34 | 30.6 | 42.5 | 5719 | 2 | AC021328 | AC021328 Homo sapi |
| 35 | 30.6 | 42.5 | 204412 | 2 | AF327569 | AF327569 Mus muscu |
| 36 | 30.6 | 42.2 | 181061 | 2 | AC021446 | AC021446 Mus muscu |
| 37 | 30.4 | 42.2 | 181061 | 2 | AC073936 | AC073936 Mus muscu |
| 38 | 30.2 | 41.9 | 143894 | 33 | AC069002 | AC069002 Homo sapi |
| 39 | 30.2 | 41.9 | 156148 | 2 | AL591038 | AL591038 Homo sapi |
| 40 | 30.2 | 41.9 | 182486 | 2 | AC03209 | AC03209 Homo sapi |
| 41 | 30 | 41.7 | 5210 | 9 | AB023137 | AB023137 Homo sapi |
| 42 | 30 | 41.7 | 7525 | 9 | AJ303079 | AJ303079 Homo sapi |
| 43 | 30 | 41.7 | 83001 | 2 | AC015632 | AC015632 Homo sapi |
| 44 | 30 | 41.7 | 89813 | 2 | AC023660 | AC023660 Homo sapi |
| 45 | 30 | 41.7 | 110000 | 2 | AL359978_0 | AL359978 Homo sapi |

ALIGNMENTS

RESULT 1
LOCUS HUMEGFRG 560 bp DNA PRI 07-NOV-1994
DEFINITION Human epidermal growth factor receptor (EGFR) gene, exon 1.
ACCESSION M11234
VERSION M11234.1 GI:181981
KEYWORDS epidermal growth factor receptor.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 560)
AUTHORS Ishii,S., Xu,Y.H., Stratton,R.H., Roe,B.A., Merlino,G.T. and
Pastan,I.
TITLE Characterization and sequence of the promoter region of the human
epidermal growth factor receptor gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 82 (15), 4920-4924 (1985)
MEDLINE 85270438
FEATURES Location/Qualifiers
SOURCE 1..560

| | |
|-----------------|---|
| RESULT | 2 |
| HSEGF1 | |
| LOCUS | HSEGFR1 1347 bp DNA PRI 24-FEB-1999 |
| DEFINITION | Human A431 EGF receptor gene 5'-flank and exon 1. |
| ACCESSION | X06370 |
| VERSION | X06370.1 GI:31118 |
| KEYWORDS | epidermal growth factor receptor; erbB proto-oncogene; growth factor receptor; proto-oncogene; receptor. |
| SOURCE | human. |
| ORGANISM | Homo sapiens |
| REFERENCE | Eukaryota: Metazoa; Chordata; Craniata; Vertebrata: Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae: Homo. |
| AUTHORS | 1 (bases 1 to 1347) Haley,C., Whittle,N., Bennet,P., Kinchington,D., Ulirsch,A. and Waterfield,M. |
| TITLE | The human EGF receptor gene: structure of the 110 kb locus and identification of sequences regulating its transcription |
| JOURNAL MEDLINE | Oncogene Res. 1 (4), 375-396 (1987) |
| FEATURES | 88217333 |
| source | Location/Qualifiers 1..1347 |

```

/codon_start=1
/protein_id="CAA29668.1"
/db_xref="GI:31119"
/db_xref="SWISS-PROT:P00533"
/translation="MRPSTAGALLALALACPASRALEEKK"
1190. >1347
/number=1
misc_feature 1216..1221
    /note="pot. Sp1 binding site"
misc_feature 1248..1251
    /note="pot. Sp1 binding site"
misc_feature 1286..1291
    /note="pot. Sp1 binding site"
BASE COUNT 210 a 508 c 398 g 231 t
ORIGIN

```

```

Query Match          100.0%: Score 72: DB 9: Length 1347:
Best Local Similarity 100.0%: Pred. No. 3e-07:
Matches 72: Conservative 0: Mismatches 0: Indels 0: Gaps 0

Oy 1 atgcgacctccgcggagcgcgcggcgagcgtctctgtgcgtctgcgtctgcgcg 60
      |||
Db 1102 ATGGCAGCCCTCCGGGAGCGCGCGGACCGCTCTCGGCGCTGCGCTGCGCCG 1161
      |||

Oy 61 gcgagtcgggct 72
      |||

Db 1162 GCGAGTCGGGCT 1173

RESULT 3
HSU48722 HSU48722 1593 bp mRNA PRI 05-NOV-1996
LOCUS
DEFINITION Human epidermal growth factor receptor precursor (EGFR) mRNA,
complete cds.
ACCESSION U48722
VERSION U48722.1 GI:1628549
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1593)
AUTHORS Reiter,J.L. and Mainhe,N.J.
TITLE A 1.8 kb alternative transcript from the human epidermal growth
factor receptor gene encodes a truncated form of the receptor
Nucleic Acids Res. 24 (20), 4050-4056 (1996)
JOURNAL MEDLINE
REFERENCE 2 (bases 1 to 1593)
AUTHORS Reiter,J.L.
TITLE Direct Submission
JOURNAL Submitted (07-FEB-1996) Jil L. Reiter, Biochemistry and Molecular

```

FEATURES Biology, Mayo Clinic, 200 First St. SW, Rochester, MN 55905, USA
 Source Location/Qualifiers
 1..1593
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7p11.2-p12"
 /tissue_type="placenta"
 1..1456
 /gene="EGFR"
 1..244
 /gene="EGFR"
 245..1462
 /note="ErbB1-S"
 /codon_start=1
 /product="epidermal growth factor receptor precursor"
 /protein_id="A50802.1"
 /db_xref="GI:1628550"
 /translation="MRPSGTAGAAIALAALCPASRALEKKVCGTSNKLTLQLE
 EDHPLSLQRMENNCVAVLGNLETTVQRYNDLSFKTTQEVAGVYLALNTVERIPLE
 NLQITRGNMYENSVALAVLSYNDANKTLKELPRNLQELHGAVRSPNALCNE
 SIQWRDIVSDFLSNMSPQNHLSGQKDEPCSCNGSCWAGENCORLTKITCAO
 CSGRCRSPSCDCHNOCAGCTGPRESDLCVRKRPDEATCKOTCPMLMINTTQ
 MDVNEKYSFEGATKCPKPRNVVYDHGSCVRAAGASYLEMEEDGVRKCKGCPK
 KCVNGISIGFPKDSLSTNATNTHKNCISISGDLHILPVAFRDSFTHTPDPQEL
 DILKTVKEITGLS"
 245..316
 /gene="EGFR"
 317..1459
 /product="epidermal growth factor receptor"
 626..634
 /gene="EGFR"
 /note="encodes Asn-linked glycosylation site"
 767..775
 /gene="EGFR"
 /note="encodes Asn-linked glycosylation site"
 830..838
 /gene="EGFR"
 /note="encodes Asn-linked glycosylation site"
 1055..1063
 /gene="EGFR"
 /note="encodes Asn-linked glycosylation site"
 1298..1306
 /gene="EGFR"
 /note="encodes Asn-linked glycosylation site"
 1325..1333
 /gene="EGFR"
 /note="encodes Asn-linked glycosylation site"
 1452..1456
 /note="encodes Asn-linked glycosylation site"
 /gene="EGFR"
 /note="5' splice donor site"
 1463..1569
 polyA_signal
 1537..1542
 polyA_site
 1569
 BASE COUNT 406 a 445 c 424 g 318 t
 ORIGIN

Query Match 100.0%; Score 72; DB 9; Length 1593;
 Best Local Similarity 100.0%; Pred. No. 2.8e-07;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgcgacctccggagccgagccgagcctcctgcgagctgctgctgctgccc 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 245 ATCGACCTCCGGACGCGCGGCGGCGACGCTCTCGCGCTGCGCTCGCCG 304
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 61 gcgagtcgggct 72
 ||||||||||||
 Db 305 GCGAGTCGGGCT 316
 ||||||||
 RESULT 4
 180039

LOCUS 180039 1868 bp DNA PAT 10-JUN-1998
 DEFINITION Sequence 1 from patent US 5708156.
 ACCESSION 180039
 VERSION 180039.1 GI:3208329
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1868)
 AUTHORS Ilekis,J.V.
 TITLE Epidermal growth factor receptor-like gene product and its uses
 JOURNAL Patent: US 5708156-A 1 13-JUN-1998;
 FEATURES
 source Location/Qualifiers
 1..1868
 /organism="unknown"
 BASE COUNT 449 a 554 c 494 g 371 t
 ORIGIN

Query Match 100.0%; Score 72; DB 6; Length 1868;
 Best Local Similarity 100.0%; Pred. No. 2.7e-07;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgcgacctccggagccgagccgagcctcctgcgagctgctgctgctgccc 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 504 ATCGACCTCCGGACGCGCGGCGGCGACGCTCTCGCGCTGCGCTCGCCG 563
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 61 gcgagtcgggct 72
 ||||||||||||
 Db 564 GCGAGTCGGGCT 575
 ||||||||

RESULT 5
 HS095089
 LOCUS HS095089 1868 bp mRNA PRI 26-APR-1997
 DEFINITION Human truncated epidermal growth factor receptor-like protein
 precursor mRNA, complete cds.
 ACCESSION U95089
 VERSION U95089.1 GI:2051984
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1868)
 AUTHORS Ilekis,J.V., Stark,B.C. and Scoccia,B.
 TITLE Possible role of variant RNA transcripts in the regulation of
 epidermal growth factor receptor expression in human placenta
 Mol. Reprod. Dev. 41 (2), 149-156 (1995)
 JOURNAL 95382957
 MEDLINE 2 (bases 1 to 1868)
 REFERENCE Ilekis,J.V., Gartli,J., Niederberger C. and Scoccia,B.
 AUTHORS Expression of a truncated Epidermal Growth Factor Receptor-like
 protein (rEGFR) in Ovarian Cancer
 Gyn. Onc. 65 (1997) In press
 JOURNAL 3 (bases 1 to 1868)
 REFERENCE Ilekis,J.V.
 AUTHORS Direct Submission
 TITLE Submitted (25-MAR-1997) Obstetrics & Gynecology, University of
 Illinois at Chicago, 820 S. Wood Street, Chicago, IL 60612, USA
 COMMENT please also see Ilekis, J. in RK Miller, HA Thiede (eds):
 "Molecular Biology and Cell Regulation of the Placenta." New York:
 Plenum, 103-117, 1991.
 FEATURES
 source Location/Qualifiers
 1..1868
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="Placenta"
 504..575
 504..1721
 /note="rEGFR: the calculated molecular weight of the
 nascent protein is approximately 42 kD (includes signal
 peptide at amino acid positions 1-24); the molecular

weight of the mature protein is approximately 80 kD:
secreted growth factor receptor"

/codon_start=1
/product="truncated epidermal growth factor receptor-like
protein precursor"

/protein_id="AA53063.1"

/db_xref="GI:2051985"

/translation="MRPSGAGALLALALCPASRALERKVCQGTGSKTLQGLTF
EDHFLSLQRMFNCEVAVLGNLFTYVQRYVDSPLKTQEVAGVLLALNTVERIPLE
NLOJIRGNMYENSVALAVLSYNDANKTGLELPMRNLOELHGVAVRSPNNPALCNE
SISQWRDIVSDFLSNMSDFONHLGSCQCDPSCNCGACGEGCOKLTKILCAQ
CSGRCKSPDCCHNOCAAGCTGPRSCDLCRKREDACTCKDTCPLMYNTTYO
MDVPEKYSFGATCVKCRPNVYVDHSCVRACGASVEMEDGVKCKGCPGR
KVCNGICIGEKDLSLNATNIRKFNKNSISGDLILPVAFRDPSFTHPRLDPQEL
DLTKVKEITGSL"

576..1718

/product="truncated epidermal growth factor receptor-like
protein"

/db_xref="GI:2051985"

/translation="MRPSGAGALLALALCPASRALERKVCQGTGSKTLQGLTF
EDHFLSLQRMFNCEVAVLGNLFTYVQRYVDSPLKTQEVAGVLLALNTVERIPLE
NLOJIRGNMYENSVALAVLSYNDANKTGLELPMRNLOELHGVAVRSPNNPALCNE
SISQWRDIVSDFLSNMSDFONHLGSCQCDPSCNCGACGEGCOKLTKILCAQ
CSGRCKSPDCCHNOCAAGCTGPRSCDLCRKREDACTCKDTCPLMYNTTYO
MDVPEKYSFGATCVKCRPNVYVDHSCVRACGASVEMEDGVKCKGCPGR
KVCNGICIGEKDLSLNATNIRKFNKNSISGDLILPVAFRDPSFTHPRLDPQEL
DLTKVKEITGSL"

576..1718

/product="truncated epidermal growth factor receptor-like
protein"

/db_xref="GI:2051985"

/translation="MRPSGAGALLALALCPASRALERKVCQGTGSKTLQGLTF
EDHFLSLQRMFNCEVAVLGNLFTYVQRYVDSPLKTQEVAGVLLALNTVERIPLE
NLOJIRGNMYENSVALAVLSYNDANKTGLELPMRNLOELHGVAVRSPNNPALCNE
SISQWRDIVSDFLSNMSDFONHLGSCQCDPSCNCGACGEGCOKLTKILCAQ
CSGRCKSPDCCHNOCAAGCTGPRSCDLCRKREDACTCKDTCPLMYNTTYO
MDVPEKYSFGATCVKCRPNVYVDHSCVRACGASVEMEDGVKCKGCPGR
KVCNGICIGEKDLSLNATNIRKFNKNSISGDLILPVAFRDPSFTHPRLDPQEL
DLTKVKEITGSL"

576..1718

/product="truncated epidermal growth factor receptor-like
protein"

/db_xref="GI:2051985"

/translation="MRPSGAGALLALALCPASRALERKVCQGTGSKTLQGLTF
EDHFLSLQRMFNCEVAVLGNLFTYVQRYVDSPLKTQEVAGVLLALNTVERIPLE
NLOJIRGNMYENSVALAVLSYNDANKTGLELPMRNLOELHGVAVRSPNNPALCNE
SISQWRDIVSDFLSNMSDFONHLGSCQCDPSCNCGACGEGCOKLTKILCAQ
CSGRCKSPDCCHNOCAAGCTGPRSCDLCRKREDACTCKDTCPLMYNTTYO
MDVPEKYSFGATCVKCRPNVYVDHSCVRACGASVEMEDGVKCKGCPGR
KVCNGICIGEKDLSLNATNIRKFNKNSISGDLILPVAFRDPSFTHPRLDPQEL
DLTKVKEITGSL"

576..1718

/product="truncated epidermal growth factor receptor-like
protein"

/db_xref="GI:2051985"

/translation="MRPSGAGALLALALCPASRALERKVCQGTGSKTLQGLTF
EDHFLSLQRMFNCEVAVLGNLFTYVQRYVDSPLKTQEVAGVLLALNTVERIPLE
NLOJIRGNMYENSVALAVLSYNDANKTGLELPMRNLOELHGVAVRSPNNPALCNE
SISQWRDIVSDFLSNMSDFONHLGSCQCDPSCNCGACGEGCOKLTKILCAQ
CSGRCKSPDCCHNOCAAGCTGPRSCDLCRKREDACTCKDTCPLMYNTTYO
MDVPEKYSFGATCVKCRPNVYVDHSCVRACGASVEMEDGVKCKGCPGR
KVCNGICIGEKDLSLNATNIRKFNKNSISGDLILPVAFRDPSFTHPRLDPQEL
DLTKVKEITGSL"

576..1718

/product="truncated epidermal growth factor receptor-like
protein"

/db_xref="GI:2051985"

/translation="MRPSGAGALLALALCPASRALERKVCQGTGSKTLQGLTF
EDHFLSLQRMFNCEVAVLGNLFTYVQRYVDSPLKTQEVAGVLLALNTVERIPLE
NLOJIRGNMYENSVALAVLSYNDANKTGLELPMRNLOELHGVAVRSPNNPALCNE
SISQWRDIVSDFLSNMSDFONHLGSCQCDPSCNCGACGEGCOKLTKILCAQ
CSGRCKSPDCCHNOCAAGCTGPRSCDLCRKREDACTCKDTCPLMYNTTYO
MDVPEKYSFGATCVKCRPNVYVDHSCVRACGASVEMEDGVKCKGCPGR
KVCNGICIGEKDLSLNATNIRKFNKNSISGDLILPVAFRDPSFTHPRLDPQEL
DLTKVKEITGSL"

576..1718

/product="truncated epidermal growth factor receptor-like
protein"

/db_xref="GI:2051985"

/translation="MRPSGAGALLALALCPASRALERKVCQGTGSKTLQGLTF
EDHFLSLQRMFNCEVAVLGNLFTYVQRYVDSPLKTQEVAGVLLALNTVERIPLE
NLOJIRGNMYENSVALAVLSYNDANKTGLELPMRNLOELHGVAVRSPNNPALCNE
SISQWRDIVSDFLSNMSDFONHLGSCQCDPSCNCGACGEGCOKLTKILCAQ
CSGRCKSPDCCHNOCAAGCTGPRSCDLCRKREDACTCKDTCPLMYNTTYO
MDVPEKYSFGATCVKCRPNVYVDHSCVRACGASVEMEDGVKCKGCPGR
KVCNGICIGEKDLSLNATNIRKFNKNSISGDLILPVAFRDPSFTHPRLDPQEL
DLTKVKEITGSL"

576..1718

/product="truncated epidermal growth factor receptor-like
protein"

/db_xref="GI:2051985"

/translation="MRPSGAGALLALALCPASRALERKVCQGTGSKTLQGLTF
EDHFLSLQRMFNCEVAVLGNLFTYVQRYVDSPLKTQEVAGVLLALNTVERIPLE
NLOJIRGNMYENSVALAVLSYNDANKTGLELPMRNLOELHGVAVRSPNNPALCNE
SISQWRDIVSDFLSNMSDFONHLGSCQCDPSCNCGACGEGCOKLTKILCAQ
CSGRCKSPDCCHNOCAAGCTGPRSCDLCRKREDACTCKDTCPLMYNTTYO
MDVPEKYSFGATCVKCRPNVYVDHSCVRACGASVEMEDGVKCKGCPGR
KVCNGICIGEKDLSLNATNIRKFNKNSISGDLILPVAFRDPSFTHPRLDPQEL
DLTKVKEITGSL"

576..1718

/product="truncated epidermal growth factor receptor-like
protein"

/db_xref="GI:2051985"

/translation="MRPSGAGALLALALCPASRALERKVCQGTGSKTLQGLTF
EDHFLSLQRMFNCEVAVLGNLFTYVQRYVDSPLKTQEVAGVLLALNTVERIPLE
NLOJIRGNMYENSVALAVLSYNDANKTGLELPMRNLOELHGVAVRSPNNPALCNE
SISQWRDIVSDFLSNMSDFONHLGSCQCDPSCNCGACGEGCOKLTKILCAQ
CSGRCKSPDCCHNOCAAGCTGPRSCDLCRKREDACTCKDTCPLMYNTTYO
MDVPEKYSFGATCVKCRPNVYVDHSCVRACGASVEMEDGVKCKGCPGR
KVCNGICIGEKDLSLNATNIRKFNKNSISGDLILPVAFRDPSFTHPRLDPQEL
DLTKVKEITGSL"

576..1718

/product="truncated epidermal growth factor receptor-like
protein"

BASE COUNT 704 a 693 c 547 t
ORIGIN Chromosome 7: 493 bp upstream of Real site.

misc_recomb

/gene="EGFR"

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="7p12-p12"

/note="EGFR(S) mRNA, G00-120-610"

/gene="EGFR"

/note="aberrant epidermal growth factor receptor"

/db_xref="taxon:9606"

/map="7p12-p12"

/note="EGFR"

/note="aberrant epidermal growth factor receptor"

/db_xref="taxon:9606"

/map="7p12-p12"

/note="EGFR"

/note="aberrant epidermal growth factor receptor"

/db_xref="taxon:9606"

/map="7p12-p12"

/note="EGFR"

/note="aberrant epidermal growth factor receptor"

/db_xref="taxon:9606"

/map="7p12-p12"

/note="EGFR"

/note="aberrant epidermal growth factor receptor"

/db_xref="taxon:9606"

/map="7p12-p12"

/note="EGFR"

/note="aberrant epidermal growth factor receptor"

/db_xref="taxon:9606"

/map="7p12-p12"

/note="EGFR"

/note="aberrant epidermal growth factor receptor"

BASE COUNT 704 a 693 c 547 t
ORIGIN Chromosome 7: 493 bp upstream of Real site.

misc_recomb

/gene="EGFR"

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="7p12-p12"

/note="EGFR"

/note="aberrant epidermal growth factor receptor"

/db_xref="taxon:9606"

/map="7p12-p12"

/note="EGFR"

/note="aberrant epidermal growth factor receptor"

/db_xref="taxon:9606"

/map="7p12-p12"

/note="EGFR"

/note="aberrant epidermal growth factor receptor"

/db_xref="taxon:9606"

/map="7p12-p12"

/note="EGFR"

/note="aberrant epidermal growth factor receptor"

/db_xref="taxon:9606"

/map="7p12-p12"

/note="EGFR"

/note="aberrant epidermal growth factor receptor"

/db_xref="taxon:9606"

/map="7p12-p12"

/note="EGFR"

/note="aberrant epidermal growth factor receptor"

/db_xref="taxon:9606"

/map="7p12-p12"

/note="EGFR"

/note="aberrant epidermal growth factor receptor"

/db_xref="taxon:9606"

ORIGIN Map position 7q12-14.

| | | | | |
|--------------------------|---------|--------------------|-----------|--------------|
| Query Match | 100.0%; | Score 72; | DB 9; | Length 3888; |
| Best Local Similarity | 100.0%; | Pred. No. 2.1e-07; | | |
| Matches 72; Conservative | 0; | Mismatches 0; | Indels 0; | Gaps 0; |

Qy 1 atgcgacccctccggaacgcgcgggacagcgtctctgacgcctcgtcgtcgtcccg 60
|||||
Db 1114 atgcgacccctccggaacgcgcgggacagcgtctctgacgcctcgtcgtcgtcccg 1173

| | | | |
|----|------|-------------|------|
| Qy | 61 | gcgagtcggt | 72 |
| | | | |
| Db | 1174 | gcgagtcggct | 1185 |

RESULT 10

| LOCUS | 3888 bp | DNA | STS | 28-SEP-1998 |
|-----------------|---------|-----|-----|-------------|
| INITIATION | | | | |
| G31681 | | | | |
| SWSS1476 | | | | |
| Eric D. Green | | | | |
| Homo sapiens | | | | |
| STS genomic, | | | | |
| sequence tagged | | | | |

| | | |
|-----------|----------|------------|
| ACCESSION | G31681 | |
| VERSION | G31681.1 | GI:1916406 |

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 3888)

Bouffard, G.G., Iyer, L.M., Idol, J.R., Braden, V.V., Cunningham, A.F.
REFERENCE
AUTHORS

TITLE A collection of 1814 human chromosome 7-specific STSS
JOURNAL Genome Res. 7 (1), 59-64 (1997)
ISSN 1043-7548

REFERENCE 2 (bases 1 to 3888)

| TITLE | HUMAN CHROMOSOME 7 STS (1997) |
|-------|-------------------------------|
| | |

COMMENT:

Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892

Fax: 3014024735

Primer A: GTTGAAGAAATTTGAGCCCAACC

Size: 115
Profile:

| | |
|-----------------|---------------------------------|
| Presoak: | 0 degrees C for 0.00 minute(s) |
| Denaturation: | 92 degrees C for 0.17 minute(s) |
| Annealing: | 60 degrees C for 1.00 minute(s) |
| Polymerization: | 72 degrees C for 1.00 minute(s) |
| PCR Cycles: | 35 |

Protocol:

| | |
|-----------------|---------------------|
| Template: | 30-100 ng |
| Primer: | each 1 μ M |
| dNTPs: | each 200 μ M |
| Taq Polymerase: | 0.05 units/ μ l |
| Total Vol: | 10 μ l |

Buffer:

MgCl₂: 1.5 mM
KCl: 100 mM
Tris-HCl: 10 mM
NH₄Cl: 5 mM
pH: 8.6

This STS has been incorporated into the NHGRI chromosome 7 physical map, but was developed by another investigator. See Genbank record: M38425 For additional information about the NHGRI chromosome 7 mapping project, see <http://www.nhgri.nih.gov/DIR/STB/CHR7>. Also see Genomics 11:548-64 (1991) [MIMID=92128037].

FEATURES

Source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7"
/clone_lib="Eric D. Green"
1..3888
gene

```

STS

primer_bind

```

primer_bind complement(2437..2456)
BASE COUNT 762 a 1219 c 1101 g 806 t
ORIGIN

```

ORIGIN

| | | | | |
|-----------------------|---------|--------------------|--------|--------------------------------|
| Query Match | 100.0%; | Score 72; | DB 11; | Length 3888; |
| Best Local Similarity | 100.0%; | Pred. No. 2.1e-07; | | |
| Matches | 72; | Conservative | 0; | Mismatches 0; Indels 0; Gaps 0 |

Qy 1 atcgacaccctccgagacgagcgagcctctgagcctgacccg 60
|||||
Db 1114 atggacccctccggacgacggcgagcgcctgagcgtgacgctgacccg 1173

| | | | |
|----|------|--------------|------|
| Oy | 61 | gcgagtcgagct | 72 |
| | | | |
| Db | 1174 | gcgagtcgagct | 1185 |

RESULT 11
AR086089
AR086089
EE22 1-
DVA
DATE 07-SEP-2000

DEFINITION
ACCESSION

| | | |
|----------|------------|-------------|
| VERSION | AR086089.1 | GI:10012855 |
| KEYWORDS | | |
| SOURCE | Unknown. | |

ПРОПРАВНО

REFERENCE
 1. (cases 1 to 3324)
 AUTHORS
 Kling, C. Richter, Kraus, M. H. and Aaronson, S. A.
 TITLE
 erbB-2 gene segments, probes, recombinant DNA and kits for
 detection
 Patent, US 5,090,553-A 2 16-NOV-1990.

FEATURES

| | | | | | | | | |
|------------|---------------------|---|------|---|------|---|------|---|
| BASE COUNT | 1472 | a | 1484 | C | 1337 | g | 1239 | t |
| ORIGIN | /organism="unknown" | | | | | | | |

ORIGIN

| | | | | |
|-----------------------|---------|--------------------|--------|---------------|
| Query Match | 100.0%; | Score 72; | DB 6; | length 5532; |
| Best Local Similarity | 100.0%; | Pred. No. 1.8e-07; | | |
| Matches | 72; | Conservative | 0; | Mismatches 0; |
| | | | Indels | 0; |
| | | | Gaps | 0; |

Qy 1 atgcacccctccggagcagccgggacagcgcctctcgtcgtcgcctcgcgcg 60
|||||
Db 187 ATGCACCCCTCCGGGACGGCGGGGACAGCTCTGGCGCTGCGCTGCGCTGCGCG 246

| | | | |
|----|-----|--------------|-----|
| Qy | 61 | gcgagtcgagct | 72 |
| | | | |
| Db | 247 | gcgagtcgagct | 258 |

RESULT 12
HSECFPRE

REFERENCE 5 (bases 1 to 177998)

AUTHORS

Waterston, R.

TITLE

Direct Submission

COMMENT

Submitted (30-SEP-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Sep 28, 1999 this sequence version replaced gi:4662682.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu/gsc

Contact: sapientewatson.wustl.edu

----- Summary Statistics

Center project name: R_DJ1091E12

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see
 http://www.nhgri.nih.gov/DIR/GRB/CHR7, send
 mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:

This clone was derived from human PAC library RPCI-5, prepared by Pterer de Jong and coworkers at the Roswell Park Cancer Institute (http://bacpac.med.buffalo.edu) using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.
 The clone may be obtained either from Genome Systems, Inc. (http://www.genomesystems.com) or Research Genetics, Inc. (http://www.resgen.com); or from Pterer de Jong.
 VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:

The clone RPS-1091E12 from base position 192688 to 192794 is an approximate representation of a tandem repeat. There are approximately 800 bases of this tandem missing according to the restriction digests: ecorv band size real at 5047, inslilco at 4288; hindiii band size real at 10980, inslilco at 10434; bamhi band size real at 5010, inslilco at 4180.

The clone sequenced to the left is RP4-791C19, 200 bp overlap. Actual start of this clone is at base position 80354 of RP4-791C19; actual end is at base position 177998 of RPS-1091E12.

FEATURES

source

1..177998

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="7"

/map="7p11.2-p21"

/clone="RPS-1091E12"

/clone_1bp="RPCI-5"

1..2564

/rpt_family="L1"

repeat_region

2565..2617

/rpt_family="Alu"

repeat_region

2744..3029

/rpt_family="Alu"

repeat_region

3037..3132

/rpt_family="L1"

repeat_region

3138..3497

/rpt_family="MnLR"

repeat_region

3586..4083

/rpt_family="L1"

repeat_region

4084..4302

/rpt_family="Alu"

repeat_region

4303..4398

/rpt_family="L1"

repeat_region

4417..4476

/rpt_family="MnLR"

repeat_region

4729..5355

/rpt_family="L1"

misc_feature

8556..10519

repeat_region

9010..9102

/note="CpG island (MGC=72.4, o/e=0.90, #CpGs=225)"

misc_feature

9109..9472

/rpt_family="CT-rich"

misc_feature

9109..9472

/note="similar to Mus musculus EST A1787580 (NID:g5335296)

repeat_region

9225..9246

/rpt_family="GC-rich"

misc_feature

9231..9472

/note="match to EST A1554567 (NID:g4486930) t25f08.x1"

misc_feature

9231..9472

/note="match to EST A1740805 (NID:g5109093) wg24d08.x1"

misc_feature

9238..9472

/note="match to EST AA934507 (NID:g3091664) o050d06.s1"

misc_feature

9240..9472

/note="match to EST A1935313 (NID:g5674183) wp16g07.x1"

misc_feature

9258..9472

/note="match to EST A1858079 (NID:g5511695) wj70b02.x1"

misc_feature

9261..9477

/note="similar to EST A1263609 (NID:g3871812) qg89g02.x1"

misc_feature

9268..9472

/note="similar to EST R35665 (NID:g792566) yH90f01.s1"

repeat_region

9487..9582

/rpt_family="GC-rich"

misc_feature

10633..11006

/note="match to EST R35059 (NID:g791960) yH86d10.r1"

repeat_region

10637..10700

/rpt_family="CA)n"

misc_feature

10668..10845

/note="match to EST AA768080 (NID:g2819095) nv96h04.s1"

repeat_region

10949..11138

/rpt_family="MERL-type"

misc_feature

11173..11567

/note="match to EST R77147 (NID:g851779) y164h11.s1"

repeat_region

11895..12057

/rpt_family="MTR"

repeat_region

13289..13321

/rpt_family="TG)n"

repeat_region

13475..13776

/rpt_family="Alu"

repeat_region

15231..15735

/rpt_family="L2"

[illegible]

| REFERENCE | TITLE | JOURNAL | MEDLINE | REFERENCE | FEATURES | SOURCE |
|---|-------|---------|---------|-----------|----------|--------|
| 97078686 | | | | | | |
| 4 (bases 192735 to 195366) | | | | | | |
| Eley,G., Frederick,L., Wang,X.Y., Smith,D.I. and James,C.D. | | | | | | |
| 3' end structure and rearrangements of EGFR in glioblastomas | | | | | | |
| Genes Chromosomes Cancer 23 (3), 248-254 (1998) | | | | | | |
| 99005167 | | | | | | |
| 5 (bases 1 to 197496) | | | | | | |
| Reiter,J.L., Threadgill,D.W., Eley,G.D., Strunk,K.E., | | | | | | |
| Danielson,A.J., Schell,Sinclair,C., Pearzall,R.S., Green,P.J., | | | | | | |
| Yee,D., Lampland,A.L., Balasubramanian,S., Crossley,T.D., | | | | | | |
| Magnuson,T.R., James,C.D. and Mithye,N.J. | | | | | | |
| Comparative genomic sequence analysis and isolation of human and mouse alternative <i>egfr</i> transcripts encoding truncated receptor isoforms | | | | | | |
| Genomics 71 (1), 1-20 (2001) | | | | | | |
| 6 (bases 1 to 197496) | | | | | | |
| Reiter,J.L. and Eley,G.D. | | | | | | |
| Direct Submission | | | | | | |
| Submitted (12-JUL-2000) Biochemistry and Molecular Biology, Mayo | | | | | | |
| Clinic, 200 First St. SW, Rochester, MN 55905, USA | | | | | | |
| Part of this sequence is also presented in Genbank Accession number | | | | | | |
| AC006977, produced as a collaborative effort between the Genome | | | | | | |
| Sequencing Center at the Washington University School of Medicine | | | | | | |
| and J.L. Reiter and G.D. Eley. | | | | | | |
| Location/Qualifiers | | | | | | |
| 1..197496 | | | | | | |
| /organism="Homo sapiens" | | | | | | |
| /db_xref="taxon:9606" | | | | | | |
| /chromosome="7" | | | | | | |
| /map="7p11.2" | | | | | | |
| complement(1..2564) | | | | | | |
| /rpt_family="L1" | | | | | | |
| /rpt_type-dispersed | | | | | | |
| complement(2565..2617) | | | | | | |
| /rpt_family="Alu" | | | | | | |
| /rpt_type-dispersed | | | | | | |
| 2744..3029 | | | | | | |
| /rpt_family="Alu" | | | | | | |
| /rpt_type-dispersed | | | | | | |
| 3037..3132 | | | | | | |
| /rpt_family="L1" | | | | | | |
| /rpt_type-dispersed | | | | | | |
| 3136..3497 | | | | | | |
| /rpt_family="MLT1A1" | | | | | | |
| /rpt_type-dispersed | | | | | | |
| complement(3586..4083) | | | | | | |
| /rpt_family="L1" | | | | | | |
| /rpt_type-dispersed | | | | | | |
| 4084..4302 | | | | | | |
| /rpt_family="Alu" | | | | | | |
| /rpt_type-dispersed | | | | | | |
| complement(4303..4398) | | | | | | |
| /rpt_family="L1" | | | | | | |
| /rpt_type-dispersed | | | | | | |
| complement(4417..4726) | | | | | | |
| /rpt_family="MSMD" | | | | | | |
| /rpt_type-dispersed | | | | | | |
| complement(4729..5555) | | | | | | |
| /rpt_family="L1" | | | | | | |
| /rpt_type-dispersed | | | | | | |
| 6512..10515 | | | | | | |
| /note="CpG Island" | | | | | | |
| 8627..12178 | | | | | | |
| /note="MSA51476" | | | | | | |
| /db_xref="dbSTS:G31681" | | | | | | |
| 1247(9219..9477)..323293 | | | | | | |
| 132544..133412 | | | | | | |
| 133595 | | | | | | |
| 141459..142653 | | | | | | |
| 142771 | | | | | | |
| 144118..144254..145397 | | | | | | |
| 146053..146400 | | | | | | |
| 146766 | | | | | | |
| 146868..146939..147770 | | | | | | |
| 147860..150245 | | | | | | |
| 150443 | | | | | | |
| 151606..151738..153840 | | | | | | |

```

161282. .161320,163090. .163231,164028. .164150,
164629. .164927,171400. .171585,181826. .181981,
182873. .182948,188822. .188968,190421. .190518,
191294. .191461,191839. .191886,192626. .192734,
193367. .197441)
/feature="EGFR"
/feature="5.8 kb transcript"
/citation="11"
/product="p170 epidermal growth factor receptor"
/feature="experimental"
join(9219. .9472,132393. .132544,133412. .133595,
136713. .136847,141401. .141469,142653. .142771,
144118. .144259,145937. .146053,146640. .146766,
146866. .147056)
/feature="EGFR"
/feature="1.8 kb alternative transcript"
/product="p60 epidermal growth factor receptor"
/feature="experimental"
join(9219. .9472,132393. .132544,133412. .133595,
136713. .136847,141401. .141469,142653. .142771,
144118. .144259,145937. .146053,146640. .146766,
146866. .146939,147770. .147860,150246. .150445,
151606. .151738,153840. .153930,155387. .155544,
161282. .161320,162953. .163030)
/feature="EGFR"
/feature="aberrant 2.8 kb A431-specific transcript"
/citation="12"
/product="A431-specific p115 epidermal growth factor
receptor"
/feature="experimental"
9219. .9384
/feature="EGFR"
9219. .9472
/feature="EGFR"
/feature="number=1"
join(9219. .9472,132393. .132544,133412. .133595,
136713. .136847,141401. .141469,142653. .142771,
144118. .144259,145937. .146053,146640. .146766,
146866. .146939,147770. .147860,150246. .150445,
151606. .151738,153840. .153930,155387. .155544,
158630. .158741)
/feature="EGFR"
/feature="2.4 kb alternative transcript"
/product="truncated epidermal growth factor receptor"
/feature="experimental"
join(9219. .9472,132393. .132544,133412. .133595,
136713. .136847,141401. .141469,142653. .142771,
144118. .144259,145937. .146053,146640. .146766,
146866. .146939,147770. .147860,150246. .150445,
151606. .151738,153840. .153930,155387. .155544,
160414. .161152)
/feature="EGFR"
/feature="3 kb alternative transcript"
/product="p110 epidermal growth factor receptor"
/feature="experimental"
9219. .197441
/feature="EGFR"
/feature="ERBB1: HER1"
/feature="experimental"
join(9385. .9472,132393. .132544,133412. .133595,
136713. .136847,141401. .141469,142653. .142771,
144118. .144259,145937. .146053,146640. .146766,
146866. .146939,147770. .147860,150246. .150445,
151606. .151738,153840. .153930,155387. .155544,
161282. .161320,162953. .163007)
/feature="EGFR"
/feature="citation=12"
/feature="codon_start=1"
/feature="experimental"
/product="A431-specific p115 epidermal growth factor
receptor"
/protein_id="AA035788.1"
/db_xref="GI:11494379"

```

```

/translation="MRPSGTAGALLALIALCAPSALAEKKKCGTSTNKLTYLGTF
EHRFUSLQRMNNECVLGNLETTYVQRNYDLSFKTIQDAVYVIALNTPERIPLE
NQIITIRGMVYENSVYALVYISNDANKGLEGLEPMRILOELIGAVRFNNRLACNVE
S1QWMDIVSDFLSNMSMDONHLSGCKPSCPNCSGACGECOKLTIKICAQO
CGRCRGKSPDSCNOCAGCGPSCPCVCRPFDEATCTDTPKMLYNTYQO
MOVNPEKVSFGATCVKCPRNYYVYDHSQVACACADSYEMEDCVRRKCEGPR
KYCNCGITGIFELLIQAMPENRTDIAFENLEIIRGRKQGOFSIAVSLNITSGL
DLKTIISDGVYIISGNKLCYANTIMKKLFGTSGOKTILSRKENSCKATGOVCHA
LCSPEGCMGPEPRDCVSCRNVSGRECVCKCNLLGEPREFEVNSICDCECLPQA
NMTCTGKGPENCICCAHYIDGPRCVKTCPGVGVGNNLTVMKYADAGHVCILCHPNC
TYGCGPGLCECPNTSSYVTSHPSPFYKVMVH"
101019385. .9472,132393. .132544,133412. .133595,
136713. .136847,141401. .141469,142653. .142771,
144118. .144259,145937. .146053,146640. .146766,
146866. .146939,147770. .147860,150246. .150445,
151606. .151738,153840. .153930,155387. .155544,
158630. .158636)
/feature="EGFR"
/feature="codon_start=1"
Query Match 100.0%; Score 72; DB 9; Length 197496;
Best Local Similarity 100.0%; Pred. No. 5.3e-08;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 atggaacctccgagcgagcgagcgagcgcctcgtgcgcgtgcgcgtgcgcgcg 60
|||||
Db 9385 ATGGGACCTCTCGGAGCGGCGGCGGCGGCGCTCTCTGCGCTCTGCGCTGCGCCG 9444
|||||
Oy 61 gcgagctggagct 72
|||||
Db 9445 GCGAGCTCGGCT 9456
|||||
RESULT 15
AB025197 1858 bp DNA ROD 01-APR-1999
LOCUS
DEFINITION
AB025197
ACCESSION
AB025197.1 GI:4586455
KEYWORDS
egfr
SOURCE
Rattus norvegicus cell_line:NRK DNA.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 1858)
AUTHORS
Kondoh S.K. and Akiyama,N.
TITLE
Ret EGFR promoter from NRK cells
JOURNAL
Published Only in Database (1999) In press
REFERENCE
2 (bases 1 to 1858)
Kondoh S.K. and Akiyama,N.
Direct Submision
Submitted (19-MAR-1999) to the DDBJ/EMBL/GenBank databases. Shiniae
K. Kondoh, Kyoto Univ. Graduate School of Medicine, Dept. of
Molecular oncology; Yoshida-Konoeg-cho, Sakyo-ku, Kyoto 606-8501,
Japan (E-mail:skondoh@viru1.virus.kyoto-u.ac.jp,
Tel:81-75-751-4152, Fax:81-75-751-4159)
Location/Qualifiers
1. 1858
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/cell_line="NRK"
1775. 1858
/feature="egfr"
1775. >1858
/feature="egfr"
1775. >1858
/feature="egfr"
/feature="codon_start=1"
/feature="product="egfr"
/protein_id="BAA76391.1"
/db_xref="GI:4586456"
/translation="MRPSGTARTKLLILALACAGALEEK"
BASE COUNT 424 a 510 c 471 g 453 t

```


•
•
•
•

•

• •

XX Claim 6: page 27-28; 31pp: Japanese.
 PS
 CC This invention describes a novel rapid and highly reproducible method for
 CC diagnosis of a cell proliferative disease, comprising analyzing the
 CC extent of methylation of cytosine residue in a region participating in
 CC the chromosome DNA-related expression of a cytokine receptor gene. The
 CC method is used for the diagnosis for psoriasis, chronic rheumatoid
 CC arthritis, arteriosclerosis, blood vessel re-narrowing, diabetic
 CC retinopathy, premature neonate retinopathy or solid tumor using body
 CC fluids or tissues, particularly blood for genomic DNA extraction. This
 CC sequence represents fragment of the human cytokine receptor described in
 CC the method of the invention.
 CC
 SQ Sequence 1200 BP; 195 A; 438 C; 354 G; 213 T; 0 other;
 XX
 XX Query Match 100.0%; Score 72; DB 21; Length 1200;
 XX Best Local Similarity 100.0%; Pred. No. 7.5e-10;
 XX Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 Db 1114 atgcgacctccgagcggcgccgctcctcgtgctgctgctgctgctgccc 60
 1 atgcgacctccgagcggcgccgctcctcgtgctgctgctgctgctgccc 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1114 atgcgacctccgagcggcgccgctcctcgtgctgctgctgctgctgccc 1173
 1174 gcgagtcgggct 1185
 Oy 61 gcgagtcgggct 72
 Db 1174 gcgagtcgggct 1185
 XX
 XX RESULT 2
 XX AAS01680
 XX ID AAS01680 standard; DNA; 1663 BP.
 XX
 XX AAS01680:
 XX
 XX 18-JUL-2001 (first entry)
 XX
 XX Human epidermal growth factor receptor (EGFR) Cpg island.
 XX
 XX Human: T-type calcium channel; CACNA1G; cytosine methylation; Cpg island;
 XX cellular proliferative disorder; colorectal cancer; age related disease;
 XX apolipoprotein B; APOB; caudal type homeobox transcription factor 2;
 XX CDX2; epidermal growth factor receptor; EGFR; fibroblast-1; FBNI;
 XX G protein-coupled receptor 37; GPR37; heat shock 70KD protein 6; HSP70B;
 XX HSP66; RasGAP-related protein; IQGAP2; proteinase-activated receptor 2;
 XX PAR2; paired-like homeodomain transcription factor 2; PITX2; klotho; KL;
 XX patched A; patched B; PTCHA; PTCHB; syndecan 1; syndecan 4; SDC1; SDC4;
 XX chromosome 7p12; ds.
 XX
 XX Homo sapiens.
 XX
 XX WO200119845-A1.
 XX
 XX 22-MAR-2001.
 XX
 XX 14-SEP-2000; 2000MO-US25479.
 XX
 XX 15-SEP-1999; 99US-0398522.
 XX
 XX (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
 XX
 XX Issa J;
 XX
 XX WPI; 2001-24477/25.
 XX
 XX New nucleic acid molecule for use as a marker for screening cancer,
 XX PT comprises the coding region for a T-type calcium channel and regulatory
 XX PT sequences associated with the channel -
 XX
 XX Claim 32; Fig 4B; 125pp; English.
 XX
 XX The present sequence for epidermal growth factor receptor (EGFR)
 CC

CC Cpg island is used to study the methylation state of EGFR which
 CC maps to chromosome 7p12. The methylation state of specific regions
 CC within Cpg islands associated with a novel T-type calcium channel CACNA1G
 CC gene correlate with several cancerous phenotypes involving various tissue
 CC and cell types. Since aberrant methylation of normally unmethylated Cpg
 CC islands is often observed in immortalized and transformed cells, CACNA1G
 CC is implicated in cellular proliferative disorders e.g. leukemia,
 CC colorectal, lung, breast and other cancers. The nucleic acid coding for
 CC CACNA1G is useful as a marker for screening cancer and age related
 CC diseases. A diagnostic kit containing primers (AAS01574-AAS01623) for
 CC amplification of a Cpg-containing nucleic acid, where the primer
 CC hybridizes with a target polynucleotide sequence (AAS01627-AAS01676),
 CC can be used for detecting aberrant methylation. The Cpg island sequences
 CC (AAS01677-AAS01692) are selected from genes encoding CACNA1G,
 CC apolipoprotein B (APOB), caudal type homeobox transcription factor 2
 CC (CDX2), epidermal growth factor receptor (EGFR), fibroblast-1 (FBNI),
 CC G protein-coupled receptor 37 (GPR37), heat shock 70KD protein 6
 CC (HSP70B; HSP66), RasGAP-related protein (IQGAP2), klotho (KL),
 CC proteinase-activated receptor 2 (PAR2), paired-like homeodomain
 CC transcription factor 2 (PITX2), patched A and B (PTCHA; PTCHB) and
 CC syndecan 1 and 4 (SDC1; SDC4) or a MINT31 sequence.
 CC
 SQ Sequence 1663 BP; 199 A; 684 C; 543 G; 237 T; 0 other;
 XX
 XX Query Match 100.0%; Score 72; DB 22; Length 1663;
 XX Best Local Similarity 100.0%; Pred. NO. 7.4e-10;
 XX Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 Oy 1 atgcgacctccgagcggcgccgctcctcgtgctgctgctgctgctgccc 60
 546 atgcgacctccgagcggcgccgctcctcgtgctgctgctgctgctgccc 605
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 61 gcgagtcgggct 72
 Db 606 gcgagtcgggct 617
 XX
 XX RESULT 3
 XX AAV06408
 XX ID AAV06408 standard; DNA; 1668 BP.
 XX
 XX AAV06408:
 XX
 XX 05-MAY-1998 (first entry)
 XX
 XX Epidermal growth factor receptor-like protein (TEGFR) encoding cDNA.
 XX
 XX Epidermal growth factor receptor-like protein; TEGFR; human; antibody;
 XX diagnosis; cancer; hyperproliferative disease; ss.
 XX
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX CDS 504..1721
 XX FT /*tag= a
 XX FT /product= TEGFR
 XX FT 504..575
 XX FT /*tag= b
 XX FT /note= "putative signal peptide"
 XX FT 576..1718
 XX FT /*tag= c
 XX
 XX US5708156-A.
 XX
 XX 13-JAN-1998.
 XX
 XX 31-MAY-1996; 96US-0658883.
 XX
 XX 31-MAY-1996; 96US-0658883.
 XX
 XX (ILEK/) ILEKIS J V.
 XX
 XX

```

XX Ilekis JV;
XX WPI: 1998-100411/09.
DR P-PDB; AAM33737.
XX
XX Polypeptide related to epidermal growth factor receptor - and
PT corresponding DNA and antibody useful for cancer diagnosis
XX
PS Claim 3; Columns 11-16; 17pp; English.
XX
CC This CDNA encodes a novel human epidermal growth factor receptor-like
CC protein termed TEGFR. TEGFR is derived from human placenta. Fragments of
CC the TEGFR cDNA can be used as a probe in a hybridisation assay to detect
CC full length of TEGFR. An antibody can be produced using the antigen
CC comprising at least the unique C-terminal sequence of the TEGFR
CC polypeptide, optionally conjugated to an immunogenic carrier. The
CC antibody can be used in an immunoassay to detect the polypeptide. The
CC hybridisation assay and immunoassay can be used in the diagnosis of human
CC cancers. Antibodies and other inhibitors of the TEGFR polypeptide may
CC also be used to treat hyperproliferative diseases including cancer.
CC
GQ Sequence 1868 BP; 449 A; 554 C; 494 G; 371 T; 0 other;
XX
Query Match          100.0%; Score 72; DB 19; Length 1868;
Best Local Similarity 100.0%; Pred. No. 7,4e-10;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY      1 atgcgaccctcggcgagcgccgacgctccttggtggtgctgctgccg 60
        |||||||
Db       504 atgcgaccctcggcgagcgccgacgctccttggtggtgctgctgccg 563
OY      QY      61 gcgaagtcggagct 72
        |||||||
Db       564 gcgaagtcggagct 575
XX
RESULT   4
ID        AAF85332
AC        AAF85332 standard; DNA; 3633 BP.
XX
AC        AAF85332;
XX
DT        23-JUL-2001 (first entry)
DE        Nucleotide sequence of wild type EGFR1.
XX
OS        Homo sapiens.
XX
FH        Key                      Location/Qualifiers
FT        CDS                       1..3633
FT                                     /cds_tag=a
FT                                     /transl_except="(pos: 1618..1620, aa: Asn)"
FT                                     /product="EGFR1"
XX
XX MOZ00136659-A2.
XX
XX 25-MAY-2001.
XX
XX 17-NOV-2000; 2000MO-BP11474.
XX
XX 19-NOV-1999; 99US-0166594.
XX 19-NOV-1999; 99US-0444038.
XX 30-MAR-2000; 2000US-0539248.
XX
PA (NOVS ) NOVARTIS AG.
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
XX pipig SD, Veres G;

```

```

XX MPI: 2001-343832/36.
DR P-PSDB; AAB68420.
XX
PT Identifying genetically modified mammalian cells for use in e.g. bone
PT marrow transplants, comprises use of mutated protein-tyrosine kinase
PT receptor nucleic acid.
XX
PS Example 1; Fig 3; 63pp; English.
XX
CC The present sequence encodes a human epidermal growth factor receptor 1
CC (EGFR). EGFR is a protein-tyrosine kinase receptor (PTKR). The
CC specification describes a method for identifying genetically modified
CC mammalian cells. The method uses a mutated PTKR nucleic acid which
CC comprises a modification to the intracellular and extracellular domains,
CC or comprises a modification to the intracellular domain and excludes
CC any nerve growth factor receptor(s) (NGFR). The method uses mutated PTKR
CC as a cell surface marker, and is useful for identifying genetically
CC modified cells, especially immunoselection of transduced mammalian cells,
CC and for identifying mammalian cells expressing a protein of interest.
CC The genetically modified mammalian cells may be used in an autologous or
CC allogeneic setting e.g. gene therapy for bone marrow transplants, graft
CC facilitation or immune reconstitution.
XX
SQ Sequence 3633 BP; 946 A; 1007 C; 953 G; 727 T; 0 other;
XX
Query Match 100.0%; Score 72; DB 22; Length 3633;
Best Local Similarity 100.0%; Pred. No. 7; Je-10;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 atgcgacctccgagcgagcgggcgagcgtctctgtcgctgtgctgcgtctcccg 60
    |||||||
Db 1 atgcgacctccgagcgagcgggcgagcgtctctgtcgctgtgctgcgtctcccg 60
OY 61 gcgagctcgagct 72
    |||||||
Db 61 gcgagctcgagct 72
    |||||||

RESULT 5
AAC61552
ID AAC61552 standard; cDNA; 3818 BP.
XX
AC AAC61552;
XX
DT 19-FEB-2001 (first entry)
XX
DE cDNA sequence encoding an epidermal growth factor receptor.
XX
KM Glioma; brain; cancer; epidermal growth factor receptor; EGFR; mutant;
KW breast tumour; lung tumour; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 187..3817
FT FT /*tag= "a
FT FT /product= "epidermal growth factor receptor"
FT FT /note= "no termination codon given"
XX
PN US6127126-A.
XX
PD 03-OCT-2000.
XX
PF 09-MAR-1999; 99US-0264723.
XX
PR 07-JUN-1995; 95US-0479808.
PR 01-JUN-1990; 90US-0531410.
PR 08-SEP-1989; 89US-0404226.
PR 11-JUN-1992; 92US-0896509.
XX
PA (UTD0-) UNIT DUKE.

```

PA (UYJO) UNIV JOHNS HOPKINS.
 XX Bligner D, Vogelstein B;
 XX WPI: 2000-601693/57.
 DR P-PSDB: AAB19259.
 XX
 PT Diagnosing glioma in brain cell sample involves determining the
 PT presence of a nucleic acid encoding an epidermal growth factor receptor
 PT mutant protein type II -
 PS Disclosure: Fig 11A-F; 51pp; English.
 XX The specification describes a method for diagnosing glioma in a brain
 CC cell sample suspected of being cancerous. The method comprises detecting
 CC the presence of a nucleic acid encoding an epidermal growth factor
 CC receptor (EGFR) mutant protein type II. Deletions in the EGFR gene
 CC are found in many gliomas, breast tumours and lung tumours. The method
 CC is useful for diagnosing gliomas, breast tumours and lung tumours. The
 CC present sequence encodes a normal EGFR polypeptide.

Sequence 3818 BP; 972 A; 1089 C; 1012 G; 745 T; 0 other;

Query Match Best Local Similarity 100.0%; Score 72; DB 21; Length 3818;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgcgacccctccgcgagcgccgagcgctcctcgtcgtcgtcgtcgtccg 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 187 atgcgacccctccgcgagcgccgagcgctcctcgtcgtcgtcgtcgtccg 246
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 61 gcgagtcgggct 72
 ||||||||||||||||
 Db 247 gcgagtcgggct 258

RESULT 6
 ID AA23954
 XX AA23954 standard; DNA: 5532 BP.
 AC AA23954;
 XX
 DT 28-JAN-2000 (first entry)
 DE Human EGF receptor DNA.
 XX Human: V-erbB; MAC117; epidermal growth factor receptor; anticancer;
 XX cytosolic; toxin; cancer; treatment; detection; mammary carcinoma;
 XX malignant; EGF receptor; ss.
 XX Homo sapiens.
 OS
 XX
 FT Key Location/Qualifiers
 FT CDS 187..3819
 FT /*tag= a
 FT /product= "EGF receptor"

US5985553-A.
 16-NOV-1999.
 07-JUN-1995; 95US-0475035.
 21-OCT-1987; 87US-0110791.
 01-NOV-1991; 91US-0786598.
 05-MAR-1986; 86US-0836414.
 (USSH) US DEPT HEALTH & HUMAN SERVICES.
 Aaronson SA, Kraus MH, King CR.
 WPI: 2000-012777/01.

DR P-PSDB: AAY50616.
 XX Purified nucleic acid useful for detection and treatment of mammary
 PT carcinoma -
 XX
 PS Disclosure: Column 19-30; 32pp; English.
 XX

CC This invention describes a novel purified nucleic acid (I) (v-erbB
 CC related gene) specifically hybridizing to all or part of a MAC117 gene
 CC and not hybridizing to nucleic acid encoding an epidermal growth factor
 CC receptor. The product of the invention has anticancer and cytostatic
 CC activity. Antibodies to the protein encoded by (I) are conjugated to
 CC toxins and kill cancer cells expressing (I). Antibodies to the protein
 CC encoded by (I) are useful for the treatment of cancer. Fragments of (I)
 CC and the MAC117 gene are useful as probes for the detection of human
 CC mammary carcinoma or other malignancies resulting from the v-erbB related
 CC gene. This sequence encodes a human epidermal growth factor (EGF)
 CC receptor described in the invention.

Sequence 5532 BP; 1472 A; 1484 C; 1337 G; 1239 T; 0 other;

Query Match Best Local Similarity 100.0%; Score 72; DB 21; Length 5532;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgcgacccctccgcgagcgccgagcgctcctcgtcgtcgtcgtcgtccg 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 187 atgcgacccctccgcgagcgccgagcgctcctcgtcgtcgtcgtcgtccg 246
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 61 gcgagtcgggct 72
 ||||||||||||||||
 Db 247 gcgagtcgggct 258

RESULT 7
 ID AAQ43814
 XX AAQ43814 standard; DNA: 5750 BP.
 AC AAQ43814;
 XX
 DT 20-OCT-1993 (first entry)
 DE pRLD3D4 construct.
 XX
 DE Epidermal growth factor receptor truncate protein; EGF;
 KW binding sites; adsorptive agents; mammalian cell growth abnormality;
 KW detection; growth; reproduction; signal transmission; ds.
 XX
 OS Synthetic.
 OS
 XX
 FT Key Location/Qualifiers
 FT misc_feature 1077..1079
 FT /*tag= a
 FT /note= "codon ARA encodes Ile"

US5218090-A.
 08-JUN-1993.
 12-JUN-1990; 90US-0536896.
 12-JUN-1990; 90US-0536896.
 26-OCT-1990; 90US-0604728.
 (WARN) WARNER LAMBERT CO.
 Connors RW.
 WPI: 1993-196297/24.
 P-PSDB: AAR38211.
 New epidermal growth factor receptor truncate proteins - which


```

FT      /tag= d
FT      /transl_except= (pos: 1..3, aa: Met)
FT      /note= "ORF 27; encodes AAB07559"
FT      5806..12294
FT      /tag= e
FT      /note= "ORF 26; encodes AAB07560"
FT      12291..15491
FT      /tag= f
FT      /note= "ORF 25; encodes AAB07561"
FT      15488..21013
FT      /tag= g
FT      /note= "ORF 24; encodes AAB07562"
FT      21010..24666
FT      /tag= h
FT      /transl_except= (pos: 1..3, aa: Met)
FT      /note= "ORF 23; encodes AAB07563"
FT      24663..32690
FT      /tag= i
FT      /note= "ORF 22; encodes AAB07564"
FT      32893..34830
FT      /tag= j
FT      /note= "ORF 21; encodes AAB07565"
FT      34827..35804
FT      /tag= k
FT      /transl_except= (pos: 1..3, aa: Met)
FT      /note= "ORF 20; encodes AAB07566"
FT      35818..37302
FT      /tag= l
FT      /transl_except= (pos: 1..3, aa: Met)
FT      /note= "ORF 19; encodes AAB07567"
FT      37299..39215
FT      /tag= m
FT      /transl_except= (pos: 1..3, aa: Met)
FT      /note= "ORF 18; encodes AAB07568"
FT      39301..47181
FT      /tag= n
FT      /note= "ORF 17; encodes AAB07569"
FT      47178..49985
FT      /tag= o
FT      /transl_except= (pos: 1..3, aa: Met)
FT      /note= "ORF 16; encodes AAB07570"
FT      49982..51001
FT      /tag= p
FT      /transl_except= (pos: 1..3, aa: Met)
FT      /note= "ORF 15; encodes AAB07571"
FT      50998..52386
FT      /tag= q
FT      /transl_except= (pos: 1..3, aa: Met)
FT      /note= "ORF 14; encodes AAB07572"
FT      52383..52946
FT      /tag= r
FT      /note= "ORF 13; encodes AAB07573"
FT      53018..54190
FT      /tag= s
FT      /note= "ORF 12; encodes AAB07574"
FT      54187..55824
FT      /tag= t
FT      /note= "ORF 11; encodes AAB07575"
FT      55821..56093
FT      /tag= u
FT      /transl_except= (pos: 1..3, aa: Met)
FT      /note= "ORF 10; encodes AAB07576"
FT      56090..57586
FT      /tag= v
FT      /transl_except= (pos: 1..3, aa: Met)
FT      /note= "ORF 9; encodes AAB07577"
FT      57583..58857
FT      /tag= w
FT      /note= "ORF 8; encodes AAB07578"
FT
FT
XX      WO200040704-A1.
XX      13-JUL-2000.

```

```

XX      06-JAN-2000; 2000WO-US00445.
XX      P-PSDB: AAB07556, AAB07557, AAB07558, AAB07559, AAB07560, AAB07561,
XX      AAB07562, AAB07563, AAB07564, AAB07565, AAB07566, AAB07567, AAB07568,
XX      AAB07569, AAB07570, AAB07571, AAB07572, AAB07573, AAB07574, AAB07575,
XX      AAB07576, AAB07577, AAB07578.
XX      (REGC ) UNTV CALIFORNIA.
XX      Shen B, Du L, Sanchez C, Chen M, Edwards DJ.
XX      WPI: 2000-465974/40.
XX      P-PSDB: AAB07556, AAB07557, AAB07558, AAB07559, AAB07560, AAB07561,
XX      AAB07562, AAB07563, AAB07564, AAB07565, AAB07566, AAB07567, AAB07568,
XX      AAB07569, AAB07570, AAB07571, AAB07572, AAB07573, AAB07574, AAB07575,
XX      AAB07576, AAB07577, AAB07578.
XX      New bleomycin gene cluster components useful for peptide and/or
XX      polyketide metabolites, especially bleomycin, production and for
XX      chemically modifying biological molecules -
XX      Claim 8; Page 97-136; 162pp: English.
XX      The present sequence represents the BLM (Bleomycin) gene cluster,
XX      containing open reading frames (ORFs) 8-30. The proteins encoded
XX      by the gene cluster are useful for producing peptides and/or polyketide
XX      metabolites, especially bleomycin or bleomycin analogues. They are
XX      also useful for chemically modifying biological molecules to produce
XX      branched methyl groups, and for coupling amino acids and fatty
XX      acids. They may be reacted with an apo-carrier protein and coenzyme A
XX      to produce a holo-carrier protein. The BLM gene cluster or catalytic
XX      domains can be used individually or collectively to produce
XX      thiazolidine, thiazoline, dithiazoline and bithiazoline-containing
XX      microbial metabolites. The BLM gene cluster may also be used to produce
XX      sugars.
XX      Sequence 58857 BP; 7256 A; 25139 C; 19353 G; 7109 T; 0 other:
XX
XX      Query Match      40.6%; Score 29.2; DB 21; Length 58857;
XX      Best Local Similarity 69.0%; Pred. No. 25;
XX      Matches 40; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
XX
XX      QY      14 ggaacggcggggcgagcgtctcgtgcgtgcgtgcgtccgcggcgagtcgggc 71
XX      Db      10786 GGACGGCGGCGGCCCGCTCCCGGCGATGACGGCGGCGGCGGCTGCTCGGCTGGCC 10729
XX
XX      RESULT 12
XX      AAV23486/C
XX      ID      AAV23486 standard; DNA: 390 BP.
XX      AC      AAV23486;
XX      DT      24-JUL-1998 (first entry)
XX      DE      Pseudomonas xcpv secretion factor coding sequence.
XX      KW      Kinase; LipQ; LipR; lipase expression regulator; DNA binding regulator;
XX      sigma 54 promoter; secretion factor; lux-box binding element;
XX      oriV-box binding element; regulation cascade; ss.
XX      OS      Pseudomonas alcaligenes.
XX      PN      WO9806836-A2.
XX      PD      19-FEB-1998.
XX      PE      15-AUG-1997; 97WO-US14450.
XX      PR      16-AUG-1996; 96US-0699092.
XX      PA      (GEMV ) GENENCOR INT INC.

```

```

XX  Gerritise G, Quax WJ;
XX
XX  WPI; 1998-159528/14.
XX  P-PSDB; AAW53834.
XX
XX  Nucleic acids encoding proteins involved in the lipase regulation
XX  cascade from P. alcaligenes - useful for controlling production and
XX  secretion of heterologous proteins in P. alcaligenes
XX
XX  Claim 13; Page 37; 106pp; English.
XX
XX  This sequence encodes the xcpY secretion factor of Pseudomonas
XX  alcaligenes. The DNA represents a nucleic acid of the invention. The
XX  nucleic acids encode: (a) a kinase from a Pseudomonad that regulates the
XX  expression of a lipase; (b) a DNA binding regulator from a Pseudomonad
XX  that regulates the expression of a lipase; (c) a Pseudomonas alcaligenes
XX  upstream activating sequence; (d) a P. alcaligenes sigma 54 promoter that
XX  regulates the expression of a lipase; (e) a P. alcaligenes secretion
XX  factor selected from xcpP, O, R, S, T, U, V, W, X, Y and Z, and OrfV, X,
XX  Y; (f) a P. alcaligenes lux-box binding element; and (g) a OrfV-box
XX  binding element. The nucleic acids represent parts of a regulation
XX  cascade, comprising at its heart a kinase and a DNA binding regulator.
XX  These sequences can be used for the production of heterologous proteins
XX  in a host cell. The cascade also comprises secretion factors which can
XX  enhance the secretion of produced proteins.
XX
XX  Sequence 390 BP; 57 A; 127 C; 154 G; 52 T; 0 other;
XX
XX
XX  Query Match 39.7%; Score 28.6; DB 19; Length 390;
XX  Best Local Similarity 67.8%; Pred. No. 44;
XX  Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
XX
XX  2 tgcagacctccggagccggcgagcgcgtctctgtgcgtgtgctgcgtcgtcccg 60
XX  ||||| 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
XX  248 tgcacacgccacgcgcgcgcgcgcgtacgtctcctgcctcgtccgcgcgcgcgcgc 190
XX
XX  RESULT 13
XX  AAAL13900/c
XX  ID AAAL13900 standard; DNA; 390 BP.
XX
XX  AAAL13900;
XX
XX  01-AUG-2000 (first entry)
XX
XX  Pseudomonas alcaligenes XcpY nucleotide sequence SEQ ID NO:20.
XX
XX  Pseudomonas alcaligenes; expression; lipase regulation cascade;
XX  kinase; DNA binding regulator; polymerase; promoter; secretion factor;
XX  XcpP; XcpQ; XcpR; XcpS; XcpT; XcpU; XcpV; XcpW; XcpX; XcpY; XcpZ;
XX  OrfV; OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO; OrfP;
XX  OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA;
XX  OrfB; OrfC; OrfD; OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL;
XX  OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV; OrfW;
XX  OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE; OrfF; OrfG; OrfH;
XX  OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS;
XX  OrfT; OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD;
XX  OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO;
XX  OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ;
XX  OrfA; OrfB; OrfC; OrfD; OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK;
XX  OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV;
XX  OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE; OrfF; OrfG;
XX  OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR;
XX  OrfS; OrfT; OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC;
XX  OrfD; OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN;
XX  OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV; OrfW; OrfX; OrfY;
XX  OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ;
XX  OrfK; OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU;
XX  OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE; OrfF;
XX  OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ;
XX  OrfR; OrfS; OrfT; OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB;
XX  OrfC; OrfD; OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM;
XX  OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV; OrfW; OrfX;
XX  OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE; OrfF; OrfG; OrfH; OrfI;
XX  OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT;
XX  OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE;
XX  OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO; OrfP;
XX  OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA;
XX  OrfB; OrfC; OrfD; OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL;
XX  OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV; OrfW;
XX  OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE; OrfF; OrfG; OrfH;
XX  OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS;
XX  OrfT; OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD;
XX  OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO;
XX  OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ;
XX  OrfA; OrfB; OrfC; OrfD; OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK;
XX  OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV;
XX  OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE; OrfF; OrfG;
XX  OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR;
XX  OrfS; OrfT; OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC;
XX  OrfD; OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN;
XX  OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV; OrfW; OrfX; OrfY;
XX  OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ;
XX  OrfK; OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU;
XX  OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE; OrfF;
XX  OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ;
XX  OrfR; OrfS; OrfT; OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB;
XX  OrfC; OrfD; OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM;
XX  OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV; OrfW; OrfX;
XX  OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE; OrfF; OrfG; OrfH; OrfI;
XX  OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT;
XX  OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE;
XX  OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO; OrfP;
XX  OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA;
XX  OrfB; OrfC; OrfD; OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL;
XX  OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV; OrfW;
XX  OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE; OrfF; OrfG; OrfH;
XX  OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS;
XX  OrfT; OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD;
XX  OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO;
XX  OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ;
XX  OrfA; OrfB; OrfC; OrfD; OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK;
XX  OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV;
XX  OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE; OrfF; OrfG;
XX  OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN; OrfO; OrfP; OrfQ; OrfR;
XX  OrfS; OrfT; OrfU; OrfV; OrfW; OrfX; OrfY; OrfZ; OrfA; OrfB; OrfC;
XX  OrfD; OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ; OrfK; OrfL; OrfM; OrfN;
XX  OrfO; OrfP; OrfQ; OrfR; OrfS; OrfT; OrfU; OrfV; OrfW; OrfX; OrfY;
XX  OrfZ; OrfA; OrfB; OrfC; OrfD; OrfE; OrfF; OrfG; OrfH; OrfI; OrfJ;
```

[illegible]

Tue Jan 15 08:29:44 2002

us-09-715-249-1_copy_1_72.rng

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2002, 19:35:59 ; Search time 1700.56 Seconds

454.965 million cell updates/sec

Title: US-09-715-249-1_COPY_1_72

Sequence: 1 atgcgaccctccggaagc.....tctgccgcgagtcgcgt 72

Scoring table:

| | IDENTITY_NUC |
|-----------|--------------|
| Score 100 | Percent 100 |

Searched: 11351937 seqs, 5372889281 residues

total number of hits satisfying chosen parameters: 22703874

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Listing first 45 summaries

Database :

```
EST: *
1: em_estfun: *
2: em_esthum: *
3: em_estin: *
```

```

4: em_estim:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: gb_estl:*
11: gb_estr:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_iv:*
17: em_gss_pli:*
18: em_gss_pro:*
19: em_gss_rdi:*
20: em_gss_vtr:*
21: em_gss_vtr:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 72 | 100.0 | 243 | 10 | AM138053 | AM138053 UT-H-B11 |
| 2 | 72 | 100.0 | 300 | 10 | A0099322 | A0099322 AU099322 |
| 3 | 72 | 100.0 | 444 | 10 | BE327106 | BE327106 hw07b09.x |
| 4 | 72 | 100.0 | 475 | 11 | BF056131 | BF056131 7188605.x |
| 5 | 72 | 100.0 | 476 | 11 | BF058287 | BF058287 7k29a11.x |
| 6 | 72 | 100.0 | 561 | 10 | AM170373 | AM170373 xw6e06.x |
| 7 | 72 | 100.0 | 582 | 10 | A1858079 | A1858079 wj70b02.x |
| 8 | 72 | 100.0 | 645 | 10 | AM163375 | AM163375 au94d11.y |
| 9 | 72 | 100.0 | 646 | 10 | AM163038 | AM163038 au91e07.y |
| 10 | 72 | 100.0 | 678 | 10 | A1554567 | A1554567 tn25f08.x |
| 11 | 72 | 100.0 | 757 | 10 | A1935513 | A1935513 wp16g07.x |
| 12 | 72 | 100.0 | 798 | 10 | A1761585 | A1761585 wq66b11.x |

| | | | | | | |
|----|------|-------|------|----|----------|---------------------|
| 13 | 72 | 100.0 | 805 | 10 | A1740805 | A1740805 wq2408.x |
| 14 | 60.6 | 84.2 | 502 | 10 | A1763609 | A1763609 q689902.x |
| 15 | 56.8 | 78.9 | 340 | 13 | A0938957 | A0938957 N11-0011C |
| 16 | 54.6 | 75.8 | 214 | 11 | Bf112888 | Bf112888 7143009.x |
| 17 | 52.8 | 75.3 | 349 | 11 | Bf514337 | Bf514337 ut-f-h-BM1 |
| 18 | 46.4 | 64.4 | 301 | 11 | R35665 | R35665 y190601.s1 |
| 19 | 40.6 | 56.4 | 657 | 10 | A1878580 | A1878580 y120611.y |
| 20 | 40.6 | 55.4 | 740 | 11 | Bf120281 | Bf120281 6028687822 |
| 21 | 40.6 | 55.4 | 906 | 11 | Bf232802 | Bf232802 60203406 |
| 22 | 40.6 | 55.4 | 955 | 11 | Bf353473 | Bf353473 602073728 |
| 23 | 40.6 | 56.4 | 2456 | 12 | AK004911 | AK004911 Mus musc |
| 24 | 40.6 | 55.4 | 2662 | 12 | AK004883 | AK004883 Mus musc |
| 25 | 40.6 | 56.4 | 2936 | 12 | AK004944 | AK004944 Mus musc |
| 26 | 37.4 | 51.9 | 1053 | 12 | AK014017 | AK014017 Mus musc |
| 27 | 35.4 | 49.2 | 430 | 12 | Bf660321 | Bf660321 maa29d03. |
| 28 | 33 | 45.8 | 654 | 10 | Bf632171 | Bf632171 DGI BA.F1 |
| 29 | 32.6 | 45.3 | 165 | 11 | Bf903827 | Bf903827 MR1-M028 |
| 30 | 32.2 | 44.7 | 562 | 10 | Bf704513 | Bf704513 SC01_0860 |
| 31 | 31.4 | 44.6 | 728 | 10 | AL552108 | AL552108 AL552108 |
| 32 | 30.8 | 42.8 | 582 | 10 | Bf453010 | Bf453010 894562108 |
| 33 | 30.6 | 42.5 | 576 | 10 | Bf471264 | Bf471264 WHE0286_B |
| 34 | 30.6 | 42.5 | 607 | 11 | Bf255166 | Bf255166 HVSME1000 |
| 35 | 30.6 | 42.5 | 679 | 10 | AL503950 | AL503950 AL503950 |
| 36 | 30.6 | 42.5 | 679 | 10 | BE413373 | BE413373 MCG011_H0 |
| 37 | 30.6 | 42.5 | 1156 | 10 | BE421073 | BE421073 HMM005_E03 |
| 38 | 30.6 | 41.7 | 371 | 10 | AV434418 | AV434418 AV434418 |
| 39 | 30 | 41.7 | 781 | 13 | CNS01PN8 | AL14205 Anopheles |
| 40 | 29.6 | 41.7 | 955 | 11 | Bf066019 | Bf066019 HV CEP001 |
| 41 | 29.6 | 41.1 | 457 | 11 | Bf605564 | Bf605564 WHE1239.D |
| 42 | 29.6 | 41.1 | 407 | 11 | Bf785106 | Bf785106 WHE1793.L |
| 43 | 29.6 | 41.1 | 486 | 11 | Bf624156 | Bf624156 HVSME2001 |
| 44 | 29.6 | 41.1 | 515 | 13 | A2848888 | A2848888 2N0150P06 |
| 45 | 29.6 | 41.1 | 945 | 11 | Bf316964 | Bf316964 601903917 |

ALIGNMENTS

RESULT

| | | | | | |
|------------|---|--------|------|-----|-------------|
| LOCUS | AW138053 | 243 bp | mRNA | EST | 29-OCT-1999 |
| DEFINITION | UI-H-BII-abw-c-07-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone | | | | |

ACCESSION AM138053

KEYWORDS

ORGANISM

REFERENCES

AUTHORS

ТОПНАТ

COMMENT

FEATURES

ings

COMMENT

/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI-CGAP_Lu5 was prepared, and ss circles were made *in vitro*. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction.

The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo.

| | | | | |
|--------------------------|---------|--------------------|--------|-------------|
| Query Match | 100.0%; | Score 72; | DB 10; | Length 444; |
| Best Local Similarity | 100.0%; | Pred. No. 1.6e-07; | | |
| Matches 72; Conservative | 0; | Mismatches | 0; | Indels 0; |

```

27 1 atgagaccctcccgagacgacgcggaacgctccctgacgtctctgctatgacccg 60
    |||
Db 155 ATGACACCCCTCCGGAACGACGCGGCTCCTGGCGCTGCGCTGCGCTCTGCCCC 214

```

oy 61 gcgagtcgggct 72
|||||
db 215 GCGAGTCGGGCT 226

RESULT 4
BF056131
15 0000 0000

| DEFINITION |
|--|
| 7J88g95.x1 Soares_NSF_F8_9W_OT_PA_p_S1 Homo sapiens cDNA clone IMAGE:3393560 3' similar to TR:Q92795 Q92795 EPIDERMAL GROWTH FACTOR RECEPTOR PRECURSOR. [1] ; mRNA sequence. |

| | | |
|----------|------------|-------------|
| VERSION | BF056131.1 | GI:10810113 |
| KEYWORDS | EST. | |
| SOURCE | human. | |

REFERENCE

1 (bases 1 to 475)

Eubaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

| TITLE | Journal |
|---|---------|
| National Cancer Institute, Cancer Genome Anatomy Project (CCAP) | |
| Tumor Gene Index | |
| Unpublished (1997) | |

Email: cgapbs-ri@mail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

| FEATURES | location/Qualifiers |
|----------|--------------------------|
| source | 1. , 475 |
| | /organism="Homo sapiens" |

```
/clone="IMAGE:3393560"  
/clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"  
/lab_host="DH10B"
```

a modified polylinker; site_1: Not I; site_2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro.

a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the follicle

309384-310919, 323208-325855 Soares Nb2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares Nb2HF8-9W pool 1:

304776-306311, 320136-322823, 326280-326663 Soares NBH
pool 1: 723720-726407, 739080-740999 Subtraction by E
Soares and M. Fatima Bonaldo."

ORIGIN

Best Local Similarity 100.0%; Pred. No. 1.6e-07;

| | Matches | 72; Conservative | 0; Mismatches | 0; Indels | 0; Gaps |
|----|---------------------|--|---------------|-----------|---------|
| OY | 1 | atcgcacctccgagcagccgagcgagcagcgcctccctgagctgctgcgcctcgcgcg | 60 | | |
| | | | | | |
| Db | 155 | atcgcagacctccgagcagcgcctccctgagcgcctgctgcgcctgcgcctgcgcgcg | 214 | | |
| | | | | | |
| | 61 cccgagctccgagcct | 72 | | | |

Db 215 GGAGTCGGGCT 226

| | | | | |
|------------|------------|--|-----|-------------|
| BF058287 | 476 bp | mRNA | EST | 16-OCT-2000 |
| LOCUS | | | | |
| DEFINITION | 7k29a11.x1 | NCI_CGAP_Ov18 Homo sapiens cDNA clone IMAGE:3476660 3' | | |
| | similar to | IP-002705 002705 FURDERPAT GROWTH FACTOR RECEPTOR | | |

| PRECUSOR. | [1] ;, mRNA sequence. |
|-----------|-----------------------|
| ACCESSION | BF058287 |
| VERSION | BF058287.1 |
| GI: | 10812183 |

| ORGANISM | Homo sapiens |
|---|--------------|
| Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; | |
| Mammalia; Mammalia; Eumammalia; Eumammalia; Eumammalia; Homo | |

REFERENCE
1 (bases 1 to 4/6)
AUTHORS
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL COMMENT
unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cga@ps-remail.nih.gov

R. Edmeister-Buck, M.D., Ph.D., cDNA Library Preparation; M. DeLuca Soares, Ph.D., cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University

Clone distribution: MCL-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/INTL, send email to: info@image.llnl.gov

| FEATURES | High quality sequence stop: 4/5. |
|----------|----------------------------------|
| SOURCE | Location/Qualifiers |
| 1. | .476 |

```

/db_xref="taxon:9606"
/clone="IMAGE:3476660"
/clone_lib="NCI-CGAP_Ov18"

```

```

/1ab_host="DH10B (phage-resistant)"
/1ab_host="Ovary: Vector: pRT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st

```

TCCTACCAATCTGACAGTCGAGACGCGCCGCGCGACATTTTATTTTTTTTTTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors
(pharmacia), digested with Not I and cloned into the Not

BASE COUNT 96 a 146 c 141 g 93 t •
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Ronaldo.

Query Match 100.0%; Score 72; DB 11; Length 476;

| | | | | | | | | | |
|---------|-----|--|----|------------|----|--------|----|------|----|
| Matches | 72; | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
| Oy | 1 | atgcaacctccgagacgcgagcgacgcctctctgctgctgctgcgtctgccttcgccg | 60 | | | | | | |

db 155 ATGCGACCTCCGGGAGCGGCGGAGCGCTCTGGCGCTGTGGCTCGGCTGGCCG 214

QY 61 gcgagtcgggct 72

Db 215 GCGAGTCGGGCT 226

| REFERENCE | AUTHORS | TITLE | JOURNAL | COMMENT |
|-----------------------|---|--------------------------------------|-----------|-------------|
| 1 (bases 1 to 582) | NCI-CGAP | http://www.ncbi.nlm.nih.gov/ncicgap. | | |
| | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), | | | |
| | Tumor Gene Index | | | |
| | Unpublished (1997) | | | |
| | Contact: Robert Strausberg, Ph.D. | | | |
| | Email: cgapb-remail.nih.gov | | | |
| | Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. | | | |
| | Emmert-Buck, M.D., Ph.D. | | | |
| | CDNA Library Preparation: M. Bento Soares, Ph.D. | | | |
| | CDNA Library Arrayed by: Greg Lennon, Ph.D. | | | |
| | DNA Sequencing by: Washington University Genome Sequencing Center | | | |
| | Clone distribution: NCI-CGAP clone distribution information can be | | | |
| | found through the I.M.A.G.E. Consortium/LLNL at: | | | |
| | www.bio.llnl.gov/dbfp/image/image.html | | | |
| | Insert length: 1510 Std Error: 0.00 | | | |
| | Seq primer: -40bp from Glbco | | | |
| | High quality sequence stop: 465. | | | |
| | Location/Qualifiers | | | |
| | 1. 582 | | | |
| | /organism="Homo sapiens" | | | |
| | /db_xref="taxon:9606" | | | |
| | /clone="IMAGE:2408139" | | | |
| | /clone_1ib="NCI-CGAP-Lu19" | | | |
| | /tissue_type="squamous cell carcinoma, poorly | | | |
| | differentiated (4 pooled tumors, including primary and | | | |
| | metastatic)" | | | |
| | /dev_stage="adult" | | | |
| | /lab_host="DH10B (phage-resistant)" | | | |
| | /note="Organ: Lung; Vector: pTZ19-pac (Pharmacia) with a | | | |
| | modified polylinker. 1st strand cDNA was prepared from | | | |
| | pooled lung tumor tissue, and was then primed with a Not I | | | |
| | - oligo(dT) primer. Double-stranded cDNA was ligated to | | | |
| | Eco RI adaptors (Pharmacia), digested with Not I and | | | |
| | cloned into the Not I and Eco RI sites of the modified | | | |
| | pTZ19 vector. Library went through one round of | | | |
| | normalization. Library constructed by Bento Soares and M. | | | |
| | Fatima Bonaldo." | | | |
| BASE COUNT | 139 a | 161 c | 158 g | 124 t |
| ORIGIN | | | | |
| Query Match | 100.0% | Score 72: | DB 10: | Length 582: |
| Best Local Similarity | 100.0% | Pred. No. 1.6e-07: | | |
| Matches 72: | Conservative 0: | Mismatches 0: | Indels 0: | Gaps 0: |
| OY 1 | atcgacacctccggagcggcgagcgcctctgagcgtctgctgctgcgtccg 60 | | | |
| | | | | |
| DB 128 | ATGCGACCCCTCCGCGGACGCCGCGGACACCGCTCTGGCGCTGCGCTGCGCTCCG 187 | | | |
| OY 61 | gcgcagtcgcgcgt 72 | | | |
| | | | | |
| DB 188 | GCGAGTCGGGCT 199 | | | |
| RESULT 8 | | | | |
| LOCUS | AM163375 | 645 bp | mRNA | EST |
| DEFINITION | au96d11.y1 schneider fetal brain 00004 Homo sapiens cDNA clone | | | |
| | IMAGE:2483925.5 similar to gb:U03193 EPIDERMAL GROWTH FACTOR | | | |
| | RECEPTOR PRECURSOR (HUMAN);, mRNA sequence. | | | |
| ACCESSION | AM163375 | | | |
| VERSION | AM163375.1 | GI:6302408 | | |
| KEYWORDS | EST. | | | |
| SOURCE | human. | | | |
| ORGANISM | Homo sapiens | | | |
| | Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; | | | |
| | 1 (bases 1 to 645) | | | |
| REFERENCE | Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., | | | |
| AUTHORS | Kritman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, | | | |
| | J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theisling, B., | | | |

| | | | | |
|-----------------------|-----------------|--------------------|-----------|-------------|
| Query Match | 100.0%; | Score 72; | DB 10; | Length 798; |
| Best Local Similarity | 100.0%; | Pred. No. 1.6e-07; | | |
| Matches 72; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0 |

| RESULT | 13 |
|------------|---|
| LOCUS | A1740805 |
| DEFINITION | A1740805 805 bp mRNA EST 19-DEC-1999 wg4d08.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:2366031 3 Similar to gb:U03193 EPIDERMAL GROWTH FACTOR RECEPTOR PRECURSOR (HUMAN); mRNA sequence. |

ORGANISM

RECEIVED

COMMENT

FEATURES

SOURCE

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2366031"
/clone_id="Soares_NSF_F8_9W_OT_PA_P_S1"
/lab_host="DH10B"
/notice="Organ: pooled; Vector: pRTT3-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in

```

| | | | | |
|--------------------------|--------|------------|----------|-------------------|
| Query Match | 100.0% | Score 72: | DB 10: | Length 805; |
| Best Local Similarity | 100.0% | Pred. No. | 1.0e-07; | |
| Matches 72: Conservative | 0: | Mismatches | 0: | Indels 0: Gaps 0: |

Db 215 GCGAGTCGGGCT 226

| | | |
|-----------|--|------------|
| VENGION | RI200009.1 | 01.00.1012 |
| KEYWORDS | EST. | |
| SOURCE | human. | |
| ORGANISM | Homo sapiens | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi | |
| | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | |
| REFERENCE | 1 (bases 1 to 502) | |

11

FEATURES

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone IMAGE:1938578"
/clone_1lb="Soares_total_fetus_Nb2HF8_9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/notes="vector: pRT73D-Pac (Pharmacia) with a modified
polylinker: Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTTCACATCTGACGTGAGGAGCGGCCCTTAATTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pRT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "

```

BASE COUNT

| | | | | |
|-------|-------|-------|-------|----------|
| 108 a | 137 c | 139 g | 117 t | 1 others |
|-------|-------|-------|-------|----------|

